



MEMBRANE TRANSLOCASES IN A MITOCHONDRIALLY DERIVED
ORGANELLE:
A STUDY OF THE *T. VAGINALIS* HYDROGENOSOME

by

Christopher Kay, Bsc, Msci

Thesis submitted to the University of Nottingham
for the degree of Doctor of Philosophy

MEDICAL LIBRARY
QUEENS MEDICAL CENTRE

Dr Sabrina D. Dyall and Dr Ian D. Kerr, supervising

School of Biomedical Sciences,

University of Nottingham Medical School

Acknowledgements

This investigation was produced through a period which saw many people step into and out of this project, but taken together their unique ideas, energy and effort have fuelled this investigation to reach beyond its original expectations.

I am fortunate to be able to thank two supervisors, my first, Dr Sabrina Dyll, who had the difficult task of turning a freshly graduated idealist into a practical minded researcher. Secondly my sincerest thanks to my second supervisor, Dr Ian Kerr, who after Sabrina's departure adopted this peculiar student, who studied this bizarre thing, and worked tirelessly first to comprehend the essentials of this investigation, and then further to support and guide a student who had assumed the remaining *Trichomonas* research in Nottingham.

I must also thank the technicians, particularly Karen Lawler, whose miraculous capabilities, especially with respect to cloning enabled this project to cover as much ground as it did in its limited timeframe and their patience particularly when we were faced with largely unsuccessful import assays. Likewise I must also mention the students who passed a year in our lab as undergraduates, and whose time and effort advanced techniques, produced proteins, and generally contributed to the project and my own experience as researcher within a group.

Finally and more generally, I would like to thank all those who if not directly encouraged my scientific curiosity have nonetheless fostered its spirit. My family first recognized in me my potential, and I thank them for providing me with the encouragement and drive to understand the natural world and making its comprehension a passion. My partner, who throughout this project has been a steadfast pillar of support especially during the turbulent times, and whom without this project would not have been completed.

I would like to thank these, and innumerable others for their contributions, and hope that this work, represents in part fulfillment of their faith.

Abstract

Eukaryote cells are the products of a complex history of interspecies interactions, with some organelles now known to have arisen through the endosymbiosis of prokaryotic cells. Whilst these organelles 'evolved' from the original endosymbionts, their evolution has not stopped within the modern eukaryote. *Trichomonas vaginalis* is protozoan parasite, with an unusual cellular biology, this species appears to lack peroxisomes, and instead of mitochondria has divergent organelles called hydrogenosomes.

The hydrogenosomes of *Trichomonas* represent one of a growing number of highly divergent organelles, which are present in species throughout the eukaryotic kingdom. In this investigation the hydrogenosomes of *Trichomonas* are investigated with respect to their preprotein membrane translocases, a multi-membraned molecular system essential for the organelle's maintenance and biogenesis. The complexity and components of such a system are unlikely to arise duplicated from a separate organellogenesis, and thus the system's architecture is expected to indicate this organelle's descent.

To reveal the structure of this system two different practical approaches were used to determine the biology of the hydrogenosome. The first builds upon work to characterise translocase kinetics and probes the nature of the hydrogenosomal membrane translocon directly. The second explores the use of candidate translocases determined from the *Trichomonas* genome through bioinformatic analysis and their development into practical investigation through the expression of tagged proteins in transformant *T.vaginalis*. These transformants were used to visualise a population of membrane proteins *in situ* within the hydrogenosome by immunofluorescence microscopy, and further to identify their associations and interactions within the hydrogenosomal membrane using protein biochemical methods.

The data produced within this study are finally brought together to present a model for the *Trichomonas* hydrogenosomal preprotein import system, as well as the first molecular characterization of its translocase components.

Abbreviations

| | |
|-----------------------|--|
| μF | Microfarad |
| μm | Micrometer, micron |
| aa. | Amino acid |
| ATP | Adenosine TriPhosphate |
| ATPase | Adenosine triphosphatase |
| Bis-acrylamide | N'N'-methylene-bisacrylamide |
| BN | Blue Native |
| bp | Base pair |
| BSA | Bovine Serum Albumin |
| Co-IP | Co -immunoprecipitation |
| C-terminal | Carboxy terminal |
| DAPI | 4',6-diamidino-2-phenylindole |
| ddH ₂ O | Double distilled water |
| DDM | n-dodecyl beta-D-maltoside |
| DMSO | Dimethylsulphoxide |
| DNA | Deoxyribonucleic acid |
| DSP | Dithiobis[succinimidyl propionate] |
| DTT | 1,4-diothiothreitol |
| <i>E. coli</i> | <i>Escherichia coli</i> |
| <i>E. dispar</i> | <i>Entamoeba dispar</i> |
| <i>E. histolytica</i> | <i>Entamoeba histolytica</i> |
| ECL | Enzymatic chemiluminescence |
| EDTA | Ethylenediaminetetraacetic acid |
| EM | Electron Microscopy |
| ER | Endoplasmic reticulum |
| EtOH | Ethanol |
| g | Gram |
| <i>g</i> | Gravity |
| <i>G. lamblia</i> | <i>Giardia lamblia</i> |
| HA | Haemagglutinin |
| HA-tag | Haemagglutinin tag |
| HEPES | n-2-Hydroxyethylpiperazine-N'-2-ethanesulphonic acid |
| HGT | Horizontal gene transfer |
| His-tag | Poly histidine-tag |
| HRP | Horse radish peroxidase |
| IHM | Inner hydrogenosomal membrane |
| IMM | Inner mitochondrial membrane |
| IMS | Inter membrane space |
| IPTG | Isopropylthio-β-D-galactosidase |
| kb | Kilobase |
| kDa | KiloDaltons |
| LB | Luria-Bertani medium |
| LGT | Lateral gene transfer |
| <i>L. major</i> | <i>Leishmania major</i> |
| M | Molar |
| mA | Milliampere |
| MBS | 3-Maleimidobenzoyl-N-hydroxysuccinimide ester |
| MeOH | Methanol |
| min. | Minute |
| ml | Millilitre |
| mRNA | Messenger RNA |

| | |
|----------------------|---|
| MS | Mass spectrometry |
| mtHsp70 | Mitochondrial heat shock protein 70 |
| mV | Millivolts |
| mwt | Molecular weight |
| nm | Nanometre |
| <i>N. ovalis</i> | <i>Nyctotherus ovalis</i> |
| N-terminal | Amino-terminal |
| OHM | Outer hydrogenosomal membrane |
| OMM | Outer mitochondrial membrane |
| OMP | Outer membrane protein |
| ORF | Open reading frame |
| PAGE | Polyacrylamide gel electrophoresis |
| PAM | Presequence translocase-associated motor |
| PBS | Phosphate buffered saline |
| PCR | Polymerase chain reaction |
| pH | $-\log_{10}[\text{H}^+]$ |
| <i>P. falciparum</i> | <i>Plasmodium falciparum</i> |
| PI | Propidium iodide |
| PM | Plasma membrane |
| PVDF | Polyvinylidene difluoride |
| RNA | Ribonucleic acid |
| SAM | Sorting and assembly machinery |
| <i>S. cerevisiae</i> | <i>Saccharomyces cerevisiae</i> |
| SDS | Sodium dodecyl sulphate |
| <i>T. brucei</i> | <i>Trypanosoma brucei</i> |
| <i>T. foetus</i> | <i>Trichomonas foetus</i> |
| <i>T. hominis</i> | <i>Trichomonas hominis</i> |
| <i>T. tenax</i> | <i>Trichomonas tenax</i> |
| <i>T. vaginalis</i> | <i>Trichomonas vaginalis</i> |
| TBS | Tris buffered saline |
| TBST | Tris buffered saline and Tween-20 |
| TIC | Translocon at the inner envelope membrane of chloroplasts |
| TIM | Translocase of the inner membrane |
| TOC | Translocon at the outer envelope membrane of chloroplasts |
| TOM | Translocase of the outer membrane |
| Tris | Tris(hydroxymethyl)methylamine |
| Triton X-100 | Iso-octylphenoxypolyethoxyethanol |
| TV | <i>Trichomonas vaginalis</i> |
| TYM | modified Tryptone Yeast Maltose medium (Diamond's medium) |
| VDAC | Voltage dependent anion channel |
| w/v | Weight per volume |
| w/w | Weight per weight |
| WT | Wild Type |
| Ψ | Membrane potential |

Contents

Contents

- 1 INTRODUCTION.....16
 - 1.1 Preface 16
 - 1.2 Origins of the eukaryotic cell and the development of metabolic organelles..... 18
 - 1.2.1 Why does symbiosis occur? The advantages of a 'composite organism' . 19
 - 1.2.2 Endosymbiosis as a critical point in the process of eukaryogenesis..... 20
 - 1.2.3 Benefits of the mitochondrial endosymbiotic event 30
 - 1.2.4 Endosymbiont to metabolic organelle..... 32
 - 1.2.5 Development of the mitochondrial preprotein translocases 34
 - 1.2.6 Transition to mitochondrion and proteomic chimerism 41
 - 1.2.7 Diversity of eukaryotic endosymbiotically derived organelles 45
 - 1.3 *T. vaginalis* a protist with a hydrogenosome 52
 - 1.3.1 Evolutionary descent of *Trichomonas vaginalis* 53
 - 1.3.2 *Trichomonas* as a clinically significant pathogen 54
 - 1.3.3 The *Trichomonas* hydrogenosome 55
 - 1.4 Why study the preprotein import machinery in the hydrogenosomes of *T. Vaginalis*? 58
 - 1.4.1 Why *T. vaginalis*? 59
 - 1.4.2 Why the hydrogenosomes of *T. vaginalis*? 59
 - 1.4.3 The importance of the preprotein import machinery 60
 - 1.5 Objectives of this investigation..... 61
- 2 METHODS AND MATERIALS 63
 - 2.1 Chemicals and reagents..... 63
 - 2.2 Strains 63
 - 2.2.1 *T. vaginalis* strains..... 63
 - 2.2.2 Bacterial strains 63
 - 2.3 Vectors 63
 - 2.3.1 *T. vaginalis* expression vectors 63
 - 2.3.2 Bacterial expression vectors 64
 - 2.4 Primers 65
 - 2.5 Cell culture 66
 - 2.5.1 Bacterial cell culture..... 66
 - 2.5.2 *T. vaginalis* culture and maintenance 66

| | | |
|--------|--|----|
| 2.6 | DNA manipulation techniques | 66 |
| 2.6.1 | Preparation of DNA | 66 |
| 2.6.2 | Agarose gel separation of DNA..... | 67 |
| 2.6.3 | Phenol chloroform extraction | 67 |
| 2.6.4 | DNA precipitation..... | 67 |
| 2.6.5 | PCR techniques..... | 68 |
| 2.6.6 | Restriction digestion | 69 |
| 2.6.7 | Ligation of DNA fragments..... | 69 |
| 2.6.8 | Transformation of bacterial cells | 69 |
| 2.6.9 | Transformation of <i>T. vaginalis</i> cells | 69 |
| 2.7 | Protein manipulation techniques..... | 70 |
| 2.7.1 | TCA protein precipitation | 70 |
| 2.7.2 | SDS polyacrylamide gel electrophoresis | 70 |
| 2.7.3 | Blue native polyacrylamide gel electrophoresis | 71 |
| 2.7.4 | Transfer of proteins to PVDF membranes | 71 |
| 2.7.5 | Induction of protein expression in BL21..... | 71 |
| 2.7.6 | Preparation of ³⁵ S labeled protein precursors..... | 72 |
| 2.7.7 | Detection of radio-labeled proteins in SDS polyacrylamide gels | 72 |
| 2.7.8 | Protein purification by His-tag under denaturing conditions | 72 |
| 2.8 | Organelle preparation..... | 73 |
| 2.8.1 | Preparation of crude hydrogenosomes and cytosol extract..... | 73 |
| 2.8.2 | Preparation of purified hydrogenosomes..... | 73 |
| 2.9 | Hydrogenosomal manipulation techniques..... | 74 |
| 2.9.1 | Solubilisation of hydrogenosomal membranes | 74 |
| 2.9.2 | Preparation of purified hydrogenosomal membranes..... | 75 |
| 2.9.3 | Crosslinking of hydrogenosomal proteins..... | 75 |
| 2.10 | Protein import assays..... | 75 |
| 2.11 | Immunological techniques..... | 76 |
| 2.11.1 | Antibodies | 76 |
| 2.11.2 | Immunological detection of proteins transferred to PVDF membranes..... | 77 |
| 2.11.3 | Co-immuno precipitation of HA-tagged proteins | 77 |
| 2.12 | Confocal Microscopy | 78 |
| 2.12.1 | Confocal system..... | 78 |
| 2.12.2 | Preparation of <i>T. vaginalis</i> cells | 78 |
| 2.12.3 | Immunodetection of HA-tagged proteins | 78 |
| 2.12.4 | Nuclear staining with propidium iodide | 78 |
| 2.12.5 | Tracking agents..... | 78 |

| | | |
|--------|--|-----|
| 2.12.6 | Imaging software | 79 |
| 2.13 | Bioinformatic analysis tools | 79 |
| 2.13.1 | Phylogenetic analysis | 80 |
| 2.1.1 | Membrane topology | 80 |
| 3 | APPROACH TO PRACTICAL WORK..... | 82 |
| 4 | PREPROTEIN IMPORT ASSAY INVESTIGATION | 86 |
| 4.1 | Generation of <i>T. vaginalis</i> hydrogenosomal preprotein expression vectors..... | 86 |
| 4.2 | Purification of recombinant preproteins..... | 91 |
| 4.3 | Optimisation of hydrogenosome preparation..... | 94 |
| 4.4 | Import assay investigation | 96 |
| 4.5 | Crosslinking import assay..... | 99 |
| 4.6 | Attempts to isolate preprotein translocon complexes..... | 102 |
| 5 | BIOINFORMATIC ANALYSIS OF THE MITOCHONDRIAL IMPORT SYSTEM IN <i>T. VAGINALIS</i> | 104 |
| 5.1 | Elements of mitochondrial import model | 104 |
| 5.2 | Complexes of the outer membrane..... | 106 |
| 5.2.1 | Identifying <i>T. vaginalis</i> homologues to the outer membrane complexes. | 108 |
| 5.2.2 | Phylogeny of the <i>T. vaginalis</i> outer membrane protein homologues | 113 |
| 5.2.3 | Sequence analysis of the <i>T. vaginalis</i> outer membrane protein candidates. | 117 |
| 5.3 | Complexes of the Inner Membrane..... | 121 |
| 5.3.1 | Identifying <i>T. vaginalis</i> homologues to the inner membrane complexes. | 123 |
| 5.4 | The PAM motor and accessory proteins to the preprotein translocase system | 131 |
| 5.5 | Candidate preprotein import related proteins identified in <i>T. vaginalis</i> | 140 |
| 6 | MICROSCOPIC ANALYSIS OF THE ULTRASTRUCTURAL FEATURES OF <i>T. VAGINALIS</i> AND LOCALISATION OF RECOMBINANT HYDROGENOSOMAL PROTEINS..... | 143 |
| 6.1 | Generation of transformant strains expressing HA tagged candidate hydrogenosomal proteins | 143 |
| 6.2 | Preparation of fixed cells for microscopy..... | 145 |
| 6.3 | Imaging of non-transformed <i>T. vaginalis</i> C1..... | 146 |
| 6.3.1 | Confocal bright field images of fixed cells..... | 146 |
| 6.3.2 | Visualizing the nuclei of fixed <i>T. vaginalis</i> cells..... | 148 |
| 6.3.3 | Visualizing the distribution of the ER compartment using an anti-BIP antibody | 153 |
| 6.3.4 | Visualizing Golgi structures using Invitrogen molecular probes NBD C6-ceramide | 156 |
| 6.3.5 | Visualizing acidic membranous compartments with Invitrogen LysoTracker. | 159 |
| 6.4 | Localisation of HA-tagged protein within hydrogenosomes | 163 |
| 6.4.1 | Localisation of frataxin to the hydrogenosome..... | 163 |
| 6.4.2 | Localization of Tim44..... | 167 |
| 6.4.3 | Localization of Sam50 homologue Hmp43 | 168 |

| | | |
|-----------|---|------------|
| 6.4.4 | Localization of the Hup3 family of proteins, candidate Tom40 homologues | 170 |
| 6.5 | Further features of the hydrogenosome | 173 |
| 7 | BIOCHEMICAL CHARACTERIZATION OF CANDIDATE PROTEINS | 176 |
| 7.1 | Practical approach to transformant investigation | 176 |
| 7.2 | Localisation of candidate proteins in <i>T. vaginalis</i> cell fractions | 178 |
| 7.2.1 | Cellular localisation of recombinant HA-tagged frataxin | 178 |
| 7.2.2 | Localisation of wild-type and HA recombinant Hmp43 | 180 |
| 7.2.3 | Subcellular localization of the Hup3 family proteins | 181 |
| 7.2.4 | Subcellular localization of the Tim44 homologue | 183 |
| 7.3 | Hydrogenosomal membrane localisation of candidate proteins | 183 |
| 7.4 | Blue native PAGE of outer membrane hydrogenosomal candidates | 186 |
| 7.5 | Crosslinking of hydrogenosomal translocases candidates | 191 |
| 7.6 | Co-immunoprecipitation of candidate proteins and interaction partners | 196 |
| 7.7 | Attempts to identify the protein interaction partners of the β -barrel translocase candidates | 200 |
| 7.7.1 | Mass spectrometry combined with co-immunoprecipitation | 200 |
| 7.7.2 | Mass spectrometry of crosslinked co-immunoprecipitation products | 201 |
| 7.7.3 | Analysis of mass spectrometry attempts | 202 |
| 7.8 | In summary | 202 |
| 8 | DISCUSSION | 204 |
| 8.1 | Analysis of the preprotein import assays | 204 |
| 8.2 | Bioinformatic analysis of the <i>T. vaginalis</i> genome | 206 |
| 8.2.1 | Detection of β -barrel proteins in <i>T. vaginalis</i> | 206 |
| 8.2.2 | Detection of Tim17 domain containing proteins | 206 |
| 8.2.3 | General points | 207 |
| 8.3 | Microscopy findings | 208 |
| 8.3.1 | Ultrastructural characterisation of <i>T. vaginalis</i> | 208 |
| 8.3.2 | The <i>T. vaginalis</i> hydrogenosome | 208 |
| 8.3.3 | Hydrogenosomal localisation of hydrogenosomal candidates | 209 |
| 8.4 | Molecular biology characterisation | 210 |
| 8.4.1 | Candidate membrane localisation | 210 |
| 8.4.2 | Putative membrane complexes | 210 |
| 8.4.3 | Purification of hydrogenosomal protein complexes | 213 |
| 8.5 | In summary | 214 |
| 9 | REFERENCES | 217 |
| 10 | APPENDICES AND SUPPLEMENTARY DATA | 234 |
| 10.1 | Appendices for Chapter 1: Introduction | 235 |

| | | |
|---------|--|-----|
| 10.1.1 | Appendix 1 Polyphyly of mitochondrially derived organelles | 235 |
| 10.2 | Appendices for Chapter 2: Methods..... | 239 |
| 10.2.1 | Appendix 2 Suppliers of reagents used in this study..... | 239 |
| 10.2.2 | Appendix 3 pTagVag2 vector, and generation of Tv. Tim44 containing construct | 241 |
| 10.2.3 | Appendix 4 Generation of the β -barrel containing expression vectors from pTagVag2 | 242 |
| 10.2.4 | Appendix 5 Generation of constructs expressing Tv. ferredoxin and adenylate kinase | 243 |
| 10.3 | Appendices for Chapter 5: Bioinformatics | 244 |
| 10.3.1 | Appendix 6 Tom70 like proteins in <i>T. vaginalis</i> | 244 |
| 10.3.2 | Appendix 7 transmembrane prediction of Hup3 proteins | 245 |
| 10.3.3 | Appendix 8 transmembrane prediction of Hmp43 | 246 |
| 10.3.4 | Appendix 9 PAM homologues in <i>T.vaginalis</i> | 247 |
| 10.3.5 | Appendix 10 TopPred topology prediction of <i>S. cerevisiae</i> Tim17 superfamily proteins | 248 |
| 10.3.6 | Appendix 11 Detection of multiple Hsp70 homologues | 249 |
| 10.3.7 | Appendix 12 Hsp70-like protein families in <i>T. vaginalis</i> | 250 |
| 10.3.8 | Appendix 13 Identification of mtHsp70 homologues | 251 |
| 10.4 | Appendices for Chapter 6: Microscopy | 252 |
| 10.4.1 | Appendix 14 Confocal analysis of DAPI stained cells | 252 |
| 10.4.2 | Appendix 15 Confocal microscopy of PI stained <i>T. vaginalis</i> cells | 253 |
| 10.4.3 | Appendix 16 A 3D reconstruction of a Golgi apparatus..... | 253 |
| 10.4.4 | Appendix 17 Epifluorescent images of Transformant Hup3d | 254 |
| 10.4.5 | Appendix 18 3D reconstruction of frataxin transformant cells..... | 256 |
| 10.4.6 | Appendix 19 3D reconstruction of Tim44 transformant cells | 257 |
| 10.4.7 | Appendix 20 Distribution of HA- Hup3a..... | 258 |
| 10.4.8 | Appendix 21 Intracellular distribution of Hup3c..... | 259 |
| 10.4.9 | Appendix 22 Intracellular distribution of Hup3d..... | 260 |
| 10.4.10 | Appendix 23 Intracellular distribution of LysoTracker bright bodies | 261 |
| 10.4.11 | Appendix 24 measuring the dimensions of the hydrogenosome | 262 |

Table of figures

| | |
|--|----|
| Figure 1.2.1 The evolutionary processes shaping mitochondrially derived metabolic organelles | 18 |
| Figure 1.2.2 Phagotrophy model for eukaryogenesis..... | 24 |
| Figure 1.2.3 The oxygen toxicity model and microbial consortia | 27 |
| Figure 1.2.4 Endosymbiont contributions to the mitochondrial import model | 35 |

| | |
|---|-----|
| Figure 1.2.5 Establishment of host communication through the endosymbiont outer membrane | 37 |
| Figure 1.2.6 Establishment of permeome and inner membrane translocases in the proto-mitochondrial symbiont..... | 40 |
| Figure 1.2.7 The evolutionary origins for the modern mitochondrial proteome | 42 |
| Figure 1.2.8 Comparison of genomes of α -proteobacteria, descended symbionts, and organelles | 43 |
| Figure 1.2.9 The diversity of mitochondrially derived organelles in extant eukaryotes | 46 |
| Figure 1.2.10 A model for the divergence of mitochondrially derived organelles | 48 |
| Figure 3.1 Practical aspects of the two experimental approaches..... | 84 |
| Figure 3.2 Workflow methodology for exploring the preprotein import system of <i>T. vaginalis</i> | 85 |
| Figure 4.1.1 Preparation of the pTVAGAK plasmid | 89 |
| Figure 4.1.2 Preparation of the pTVAGFER plasmid..... | 89 |
| Figure 4.1.3 Inducible expression of the pTVAGAK and pTVAGFER expression constructs | 90 |
| Figure 4.2.1 Purification of the adenylate kinase preprotein..... | 91 |
| Figure 4.2.2 Purification of ^{35}S radiolabeled adenylate kinase from an <i>E. coli</i> expression system | 93 |
| Figure 4.2.3 Preprotein purification using a centrifugal size concentrator..... | 93 |
| Figure 4.3.1 The experimental strategy for the derivation of the <i>T. vaginalis</i> cell fractions | 95 |
| Figure 4.4.1 Import assay using denucleated organellar material and radiolabeled adenylate kinase | 96 |
| Figure 4.4.2 Preprotein association with purified hydrogenosomes in varying ATP conditions | 98 |
| Figure 4.5.1 Preprotein chemically crosslinked during association with hydrogenosomes results in higher molecular weight species..... | 100 |
| Figure 4.5.2 Optimization of chemical crosslinking for import assay reactions | 101 |
| Figure 5.1.1 The principle translocons of the preprotein import machinery of the mitochondrion | 105 |
| Figure 5.2.1 The TOM complex and its principle constituent proteins | 107 |
| Figure 5.2.2 SAM complex and its protein components. | 107 |
| Figure 5.2.3 Phylogenetic analysis of <i>T. vaginalis</i> β -barrel proteins with respect to other β -barrel families | 117 |
| Figure 5.2.4 Inter-relation of the Hup3 family proteins using the BLAST algorithm to align and score sequences..... | 117 |
| Figure 5.2.5 β -barrel signal sequences in <i>T. vaginalis</i> outer membrane β -barrels..... | 120 |
| Figure 5.3.1 The mitochondrial inner membrane translocases complexes, TIM22 and TIM23 | 121 |
| Figure 5.3.2 TopPred prediction of the transmembrane regions of the Tim17 domain in the Phat family proteins | 128 |
| Figure 5.3.3 Phylogeny of the <i>T. vaginalis</i> Phat proteins | 130 |
| Figure 5.4.1 The structural relation of the PAM motor, to its binding site Tim44 on the TIM23 complex | 131 |
| Figure 5.4.2 Phylogenetic analysis of the identified mtHsp70 candidates | 135 |
| Figure 5.4.3 A phylogenetic analysis of the <i>T. vaginalis</i> β -MPP subunit..... | 136 |
| Figure 5.4.4 A phylogenetic analysis of the Tim23 peripheral protein Tim44 homologue in <i>T. vaginalis</i> | 138 |

| | |
|---|-----|
| Figure 5.4.5 A phylogenetic analysis of the Tim23 peripheral protein Pam16 homologue in <i>T. vaginalis</i> . | 139 |
| Figure 5.5.1 <i>T. vaginalis</i> homologue proteins super imposed of the mitochondrial import model | 141 |
| Figure 5.5.2 Table illustrating the candidate proteins identified in the bioinformatic analysis, and the methods used in their detection. | 142 |
| 6.1.1 Generation of transformant <i>T. vaginalis</i> | 144 |
| Figure 6.3.1 Fixed cells visualized by confocal transmission | 147 |
| Figure 6.3.2 Nuclear localization in fixed <i>T. vaginalis</i> using DAPI counterstain | 149 |
| Figure 6.3.3 Propidium iodide better identifies the nucleus of <i>T. vaginalis</i> . | 150 |
| Figure 6.3.4 <i>T. vaginalis</i> cells incubated with propidium iodide and anti-HA antibody.... | 151 |
| Figure 6.3.5 Distribution of the nuclear marker propidium iodide in fixed C1 cells | 152 |
| Figure 6.3.6 Distribution of anti-BIP in fixed <i>T. vaginalis</i> cells | 154 |
| Figure 6.3.7 Distribution of BIP and PI trackers in reconstructed cells | 155 |
| Figure 6.3.8 Distribution of PI and Golgi marker in fixed cells | 157 |
| Figure 6.3.9 A 3D reconstruction of a Golgi apparatus | 158 |
| Figure 6.3.10 3D reconstruction of a Golgi apparatus and nucleus | 159 |
| Figure 6.3.11 Distribution of LysoTracker dye marking acid vesicles | 161 |
| Figure 6.3.12 A reconstruction of fixed cells showing distribution of LysoTracker marker | 162 |
| Figure 6.4.1 Intracellular distribution of frataxin | 164 |
| Figure 6.4.2 Reconstruction showing the structural arrangement of hydrogenosomes.. | 165 |
| Figure 6.4.3 The localization of the Tim44 <i>T. vaginalis</i> homologue | 167 |
| Figure 6.4.4 Reconstruction of hydrogenosomal surfaces in Tim44 | 168 |
| Figure 6.4.5 Distribution of the Sam50 homologue, Hmp43. | 169 |
| Figure 6.4.6 Intracellular distribution of Hup3b | 171 |
| Figure 6.4.7 Intracellular distribution of Hup3e | 172 |
| Figure 6.5.1 High resolution microscopy of hydrogenosomes in a Hup3c transformant. | 173 |
| Figure 6.5.2 Localisation patterns of selected cellular markers in <i>T. vaginalis</i> . | 175 |
| Figure 7.1.1 Practical approaches employed to investigate <i>T. vaginalis</i> transformants | 177 |
| Figure 7.2.1 Expression of recombinant protein in transformant cell lines | 178 |
| Figure 7.2.2 Cellular localisation of frataxin. | 179 |
| Figure 7.2.3 Localisation of wild type Hmp43 in <i>T. vaginalis</i> G3 | 180 |
| Figure 7.2.4 Localisation of Hmp43 in <i>T. vaginalis</i> C1 | 180 |
| Figure 7.2.5 Cellular localisation of Hmp43 | 181 |
| Figure 7.2.6 Cellular localization of the Hup3 family proteins | 182 |
| 7.2.7 Cellular localization of the Tim44 homologue | 183 |
| Figure 7.3.1 Derivation of hydrogenosomal membrane and soluble protein samples | 184 |
| Figure 7.3.2 localization of candidate proteins to the hydrogenosomal membrane | 185 |

| | |
|--|-----|
| Figure 7.4.1 Blue native PAGE of Hmp43 and Hup3 family proteins under different detergent conditions | 187 |
| Figure 7.4.2 BN PAGE of solubilised fractions from transformant Tim44 hydrogenosomes | 189 |
| Figure 7.4.3 Solubilisation profile for <i>T. vaginalis</i> Tim44 under different detergent conditions | 190 |
| Figure 7.5.1 Optimization of the crosslinkers MBS and DSP on HA-Hmp43 | 193 |
| Figure 7.5.2 Reversal of DSP crosslinking in Hmp43 transformant hydrogenosomes ... | 193 |
| Figure 7.5.3 Crosslinking of candidate hydrogenosomal translocase proteins..... | 195 |
| Figure 7.6.1 Co-immunoprecipitation of candidate proteins in 1% Triton X-100 | 197 |
| Figure 7.6.2 Crosslinking Co-IP of candidate proteins in 1% Triton X-100 | 199 |
| Figure 7.7.1 Extraction of co-immunoprecipitated species from selected β -barrel translocase candidates | 200 |
| Figure 8.1.1 Identification of potential preprotein interaction complexes on the hydrogenosomal surface of <i>T. vaginalis</i> | 205 |
| Figure 8.5.1 Proposed architecture for hydrogenosomal membrane translocase system based on the results of this investigation..... | 215 |
| Figure 10.3.1 Domain organization of Tom70 like proteins in <i>T. vaginalis</i> | 244 |
| Figure 10.3.2 Predicted β -barrel strand prediction in the Hup3 family proteins..... | 245 |
| Figure 10.3.3 Predicted transmembrane structure of Hmp43..... | 246 |
| Figure 10.3.4 Similarity of <i>T. vaginalis</i> and <i>S. cerevisiae</i> PAM components as visualised by CDD | 247 |
| Figure 10.3.5 Topology prediction of Tim17 superfamily proteins from <i>S. cerevisiae</i> | 248 |
| Figure 10.3.6 Distribution of sequence hits amongst <i>T. vaginalis</i> proteins homologous to <i>S. cerevisiae</i> mtHsp70..... | 249 |
| Figure 10.3.7 mtHsp70 and BIP homologues in <i>T. vaginalis</i> | 251 |
| Figure 10.4.1 Confocal analysis of DAPI stained cells..... | 252 |
| Figure 10.4.2 Confocal microscopy of PI stained <i>T. vaginalis</i> cells | 253 |
| Figure 10.4.3 A 3D reconstruction of a Golgi apparatus | 253 |
| Figure 10.4.4 Epifluorescent images of Hup3d | 254 |
| Figure 10.4.5 Epifluorescent images of the localisation of Hup3d | 255 |
| Figure 10.4.6 3D reconstruction of frataxin transformant cells | 256 |
| Figure 10.4.7 3D reconstruction of cells from the tim44 transformant..... | 257 |
| Figure 10.4.8 Distribution of HA-Hup3a..... | 258 |
| Figure 10.4.9 Intracellular distribution of Hup3c..... | 259 |
| Figure 10.4.10 Intracellular distribution of Hup3d | 260 |
| Figure 10.4.11 Localisation of Lysotracker bright bodies within 3D reconstructed cells | 261 |
| Figure 10.4.12 Microscopy analysis of a hydrogenosomal population in a Hup3 transformant | 262 |

Table of Tables

Table 2.1 HA tag position within *T. vaginalis* expression vectors64

Table 2.2 PCR primers used for vector production and sequencing.....66

Table 2.3 BigDye sequencing reaction profile68

Table 2.4 Properties of the detergents used in this study75

Table 2.5 Antibodies used within this study.....77

Table 4.1 Essential features of the pre-protein expression constructs used in this study.88

Table 5.1 BLAST results for TOM complex components109

Table 5.2 BLAST results for the SAM complex.....110

Table 5.3 β -barrel proteins identified from hydrogenosomal membrane MS.....111

Table 5.4 BLAST scores of *T. vaginalis* β -barrel proteins against named β -barrel proteins from other organisms.....113

Table 5.5 A Pfam analysis of the β -barrel candidate protein sequences showing predicted domains.118

Table 5.6 BLAST results for the TIM22 complex123

Table 5.7 BLAST results for the TIM23 Complex.....124

Table 5.8 Essential features of the Phat family of proteins.....126

Table 5.9 Predicted domains of the Phat family of proteins by Pfam.....126

Table 5.10 Inter-relation of the Phat family proteins using the BLAST algorithm to align and score sequences.....129

Table 5.11 BLAST results for the PAM Complex.....132

Table 6.1 Features of the transformant plasmids.....145

Table 7.1 Mass estimations for intermolecular species discovered in the biochemical analysis for *T. vaginalis* candidate translocases.....203

Table 8.1 Complexes determined for *T. vaginalis* β -barrel translocase candidates212

Table 8.2 *T. vaginalis* Tim44 homologue crosslinked species.....213

Table 10.1 Organisms harboring mitochondrially derived organelles.235

Table 10.2 Suppliers of selected materials used in this study240

Table 10.3 Identification of a mtHsp70 subclade from other dnaK homologues250

1 Introduction

1.1 Preface

The eukaryotic cell is the product of a remarkable series of evolutionary events. Its complexity and flexibility have given it a competitive edge in diverse environments, and has given rise to the macroscopic grandeur of the living world. These abilities were not inherent in the original pre-eukaryotic ancestor but have come about through the joining together of independent organisms (endosymbiosis). These relationships enabled these early cells to exploit new resources in their environment and adapt to an ever changing environment.

Over time these partner organisms could no longer function independently, and eventually a new kind of gestalt was formed as these symbionts became the organelles of modern eukaryotes (originally posited by L. Margulis-(Margulis 1970)). The original endosymbionts provided benefits to the host, and the emergent organelles have continued to evolve to fit the needs of the cell. The chloroplast is the product of a merger with a photosynthetic bacterium, enabling eukaryotic exploitation of light as an energy resource, an evolutionary strategy so successful that it has been multiply repeated (T. Cavalier-Smith 2002a). The mitochondrion is another organelle which has emerged from another endosymbiotic event, and over time its abilities have been streamlined to focus on efficient aerobic metabolism (Margulis 1970; Searcy 2003)

This thesis examines another endosymbiotically derived organelle, the hydrogenosome. This organelle is thought to have arisen from the same events which produced the mitochondrion (Embley et al. 2003), however at separate events in evolutionary history mitochondria have diverged and formed hydrogenosomes (Embley et al. 2003; Embley 2006). Whilst hydrogenosomes remain as metabolic organelles they have adapted to produce energy in largely anaerobic environments (Embley et al. 1995; Embley & Martin 2006). These demands have created organelles very different to typical mitochondria. Hydrogenosomes do not conduct oxidative phosphorylation (Muller 1993; Boxma et al. 2005), and may have had to acquire new enzymes to sustain metabolism in the absence of oxygen (Carlton et al. 2007; Boxma et al. 2007; Doležal et al. 2004; Rosenthal et al. 1997)

However the hydrogenosome maintains the essential features of an endosymbiotically derived organelle, its biogenesis mirrors that of the mitochondrion, duplicating through the division of pre-existing organelles (Benchimol, Johnson & de Souza 1996; Benchimol & Engelke 2003). Symbiont derived organelles exhibit genomic degradation, and develop proteomes where most material is now encoded in the nucleus (Adams & Palmer 2003). In the hydrogenosomes of *Trichomonas vaginalis* all organellar DNA is lost (Turner & Müller 1983), and so all its proteins are nuclear encoded.

A protein translocase system exists in mitochondria and chloroplasts to transfer nuclear encoded proteins translated in the cytosol across the membranes of these organelles (review- Schleiff & Becker 2011)). This import machinery evolved to allow the original eukaryotic host to take control of the endosymbiont but later as organellar DNA degenerated, increasingly to provide proteins essential for their maintenance and function (Gross & Bhattacharya 2009). This system is essential for organelle viability, and has had to address the mechanistic challenges of transporting large polypeptides across multiple membranes.

This thesis specifically examines the hydrogenosomes of a particularly divergent eukaryotic clade, and specifically one organism; *Trichomonas vaginalis*. Whilst the hydrogenosomes of some organisms are clearly recently diverged from mitochondrion as in *Nyctotherus ovalis* (de Graaf 2011), the hydrogenosomes of the *T. vaginalis* are highly diverged, and other species in its group are also amitochondriate (Aguilera et al. 2008; Lloyd et al. 2002).

This thesis takes particular interest in uncovering the protein translocase machinery necessary for sustaining the *T. vaginalis* hydrogenosome. It will use the commonalities of endosymbiont derived organellar import systems to detect and characterize the *T. vaginalis* preprotein import machinery using both bioinformatic and molecular biology techniques.

Consequently this investigation will aim to provide information about the origins, evolution and development of a divergent organelle, as well as physically examine the mechanisms and machinery essential for the biogenesis of an enigmatic organelle. This project aims to add to our information about divergent organelles, providing candidate proteins and complexes for the currently unknown translocases of the *T. vaginalis* hydrogenosome, and to enhance our knowledge of this organelle.

1.2 Origins of the eukaryotic cell and the development of metabolic organelles

The acquisitions of endosymbiotically derived organelles (organellogenesis) were pivotal moments in the development of eukaryotes (eukaryogenesis). No eukaryote so far characterised does not bear the mark of an endosymbiotic event (Embley & Martin 2006) and, as such markers from endosymbiont derived organelles are a definitive feature of eukaryotes.

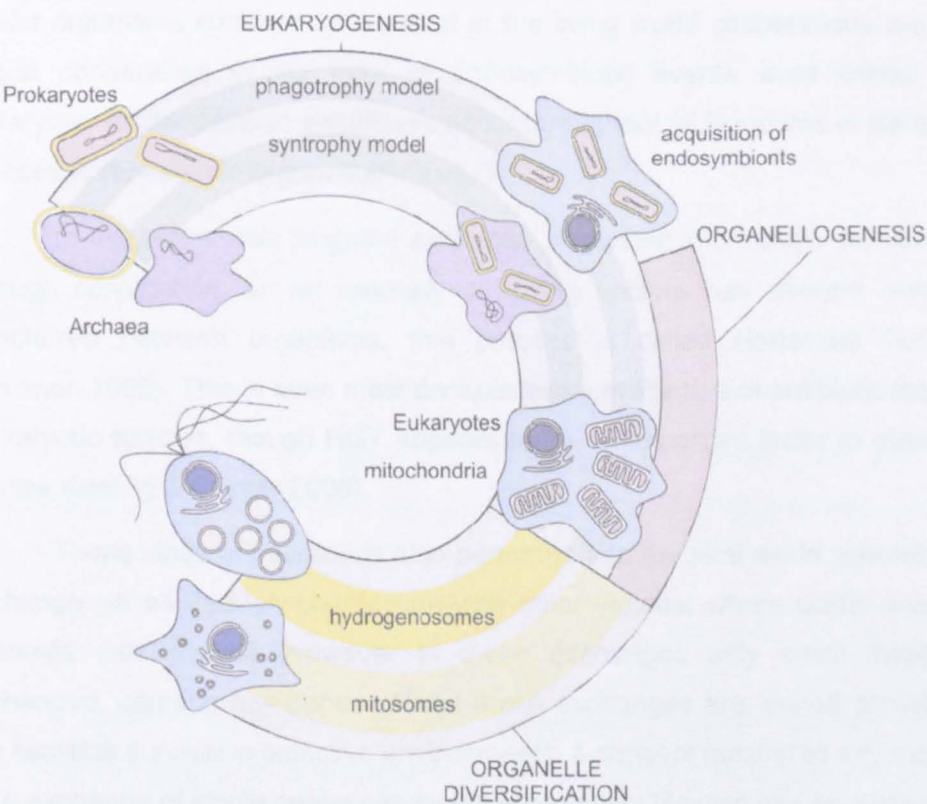


Figure 1.2.1 The evolutionary processes shaping mitochondrially derived metabolic organelles

Endosymbiotic derived organelles are a defining feature of eukaryotes and their acquisition is an essential part of the story of the evolving eukaryote. After the acquisition of the ancestral mitochondrial endosymbionts, these organisms evolved into the mitochondrion undergoing significant changes to their genome and metabolism. These organelles continued to evolve in certain eukaryotic clades giving rise to divergent organelles- hydrogenosomes and mitosomes.

This section introduces the mechanisms which brought about the endosymbiotic organelles, including the hydrogenosome. This section will look at this development in a step wise process, the forces that drive organisms to cooperate (1.2.1), and then looking at the events surrounding the endosymbiosis of the ancestral mitochondrial endosymbiont (1.2.2). Next the benefits of this endosymbiosis are briefly discussed (1.2.3) and the processes where by an endosymbiont develops into an organelle (1.2.4). Essential in the process and of specific relevance in this project are the development of membrane translocases whose origins and

development are then discussed (1.2.5). The establishment of this system brought about a new kind of informational relationship between host cell and organelle whereby the innovation of the translocase system reshaped the organelle proteome (1.2.6). Finally the processes that led to the creation of divergent organelles are introduced, and specifically address features of hydrogenosomes, both in general, and specifically in *T. vaginalis* (1.2.7).

1.2.1 Why does symbiosis occur? The advantages of a 'composite organism'

Whilst organisms compete for survival in the living world cooperations are equally ubiquitous. Whilst cooperation in the form of endosymbiotic events were critical in the developing eukaryote, similar intimate symbioses occur throughout all kingdoms of life and shed light on the processes repeated in organellogenesis.

In the prokaryotic kingdom exchange of genetic information between cells, intentionally through conjugation, or via naturally occurring vectors has allowed singular abilities to be transferred between organisms, this process is called Horizontal Gene Transfer (HGT) (Syvanen 1985). This is seen most conspicuously in the rise of antibiotic resistance in disparate prokaryotic species, though HGT appears to be an important factor in eukaryotic evolution too (review Keeling & Palmer 2008).

These kinds of interaction also penetrate into the viral world, whereby virus families can exchange genes through co-infection with other viruses, where within one organism, multiple genomes can interact. However in these exchanges only small fragments of DNA are exchanged, carrying few genes. Whilst these exchanges are limited, transfer of a single gene can facilitate survival in selective environments, a concept familiar to any molecular biologist. As such exchange of single genes can facilitate organism lifestyle and ecological niche.

More recently it has been observed that prokaryotes can form intimate interconnections with other prokaryotic cells, even between disparate species (Dubey & Ben-Yehuda 2011), forming conduits through which the exchange of substrates, and perhaps genetic material, may occur. The abilities of these communities are then the sum of their genetic diversity, multiple genomes each contributing to group survival. The benefits to survival both for individuals and the group improve with diversity; this also enables members to begin to specialize. The effect can be seen particularly clearly in some cyanobacteria, where filaments of cells share cytoplasmic connections, and individual member cells can specialize to specific tasks- such as nitrogen fixation (Elfgren 2003). These examples represent only a few of the ways in which

prokaryotic cells have developed working cooperations, and in each case each strategy brings a survival advantage.

The endosymbiotic events of the eukaryotic cell secured members of a community within a single host. This meant inter-species associations could become permanent, and a new level of integration became possible. The eukaryotic cell thus represents a community in microcosm, and like prokaryotic communities, components of a community can specialize to become efficient at specific tasks.

However the way that the eukaryotic cell lives is a radical departure from the prokaryotic communities that preceded eukaryogenesis. For the eukaryotic cell to develop there had to be a driving force behind or a benefit from eukaryogenesis which gave the evolving cells critical advantages which surpassed those of other prokaryotes. The next section discusses what role endosymbiosis had in shaping an entirely new way of living, and how processes seen here are reused in the development of the mitochondrion.

1.2.2 Endosymbiosis as a critical point in the process of eukaryogenesis

The ubiquity of endosymbiotically derived organelles and the specific impact of organellogenesis in the development of the eukaryote remains a critical unresolved argument in evolutionary biology (O'Malley 2010). However examination of the possible events that drove eukaryogenesis can explain the kinds of organelles that did emerge, and describes commonalities in their subsequent evolution.

In literature there has arisen two main theories of eukaryogenesis (ideologies reviewed in (O'Malley 2010)), one holds that the key to the whole process lay in the establishment of the first endosymbiotic relationship, where as the other suggests a route whereby the acquisition of endosymbionts was the final achievement, these key theories will be briefly described later.

In addition other theories have been raised, some discounted and others integrated into the main arguments. Of the now discounted theories some were still crucial in defining the events surrounding eukaryogenesis. Early arguments positing whether prokaryotes pre-dated eukaryotes (Forterre & Philippe 1999) and their relative ancestry have necessarily required investigations into fossil cell biology and palaeochemistry.

There is now good evidence that prokaryotic cells, existed as far back and 3.5 Gya, photosynthetic cyanobacteria as back as 2.9-7 Gya (De Marais 2000), and utilization of oxygen by these photobacteria by 2.5 Gya (Summons et al. 1999). By contrast there is great debate to the first emergence of eukaryotic cells with proposed dates as recent as 850Mya (Cavalier-

Smith 2002), but possibly as early as 2.7Gya (Hedges et al. 2004). This disparity reflects the different approaches used to try and solve this question, with evidence from molecular biology, geochemistry and palaeo-microbiology not necessarily being in close accord. However for the scope of this introduction these data indicate that eukaryotes are ultimately products of a prokaryote world. This evidence starts to set the framework for the eukaryogenesis argument.

At the time the eukaryotic cell was developing the Earth was already teeming with a fully populated prokaryotic world, with prokaryotes populating every environmental niche with great diversity in cellular chemistry. These prokaryotes are thought to be of essentially modern design. However it is at this point eukaryotic cells emerged. Some arguments on eukaryogenesis cite that an event at this time must have been critical to spur evolution (O'Malley 2010).

As already stated, oxygenic cyanobacteria are thought to have existed as early as 2.5Gya (Summons et al. 1999), and have been producing oxygen ever since. Initially in the reducing environment of primeval Earth this biochemical by-product, was sequestered by reduced iron, and organic compounds on the land and in the oceans, and in the atmosphere through UV photoreactions with organics such as methane. However over time these routes for oxygen absorption began to saturate, and oxygen began to accumulate in the environment (Holland 1994).

The effects of this transition would have reshaped the global ecology, as environmental chemistry switched from a reducing to oxidative. The upper ocean and land surface would have become aerobic environments. The establishment of oxygen in the atmosphere would have eliminated the organic aerosols and exposed the surface to UV radiation (Gross & Bhattacharya 2009) until the formation of the ozone layer, as well as wrought climactic effects produced from the elimination of methane, as for the first time we see the first global ice ages, culminating in 'snowball earth' (Hoffman et al. 1998; Kirschvink et al. 2000). These changes would have been rapid in geological terms.(Holland 1994).

Some eukaryogenesis models propose that these environmental triggers were responsible for initiating eukaryogenesis. One of the first models introducing both endosymbiosis and environmental oxygen toxicity was by Margulis (Margulis 1970), and this has model has been frequently updated as the eukaryogenesis argument has progressed (Kurland & Andersson 2000), with specific arguments developing to account other factors of eukaryogenesis such as the development of the nucleus (Martin & Koonin 2006) and sex (Gross

& Bhattacharya 2010) to be responses against an increasingly oxygenic and UV irradiated environment.

However despite the planetary scale of this change, anaerobic environments would have persisted, and thus the events of environmental oxygenation would not have entirely obliterated the preceding environment. Environmental oxygenation and photosynthetic carbon fixation might have even offered new sources of reduced carbon for anaerobic organisms; this is conspicuously seen in the formation of coal and oil deposits as well as modern peat bogs. In addition it has been suggested that trace amounts of (atmospheric) oxygen had always been present throughout Earth's history, even during the first expansion of life (Lindsay 2008).

Nevertheless the transition to an oxygen dominant environment demanded that the originally anaerobic organisms adapt to occupy the new aerobic niches. Environmental oxygenation would not be then an extreme case of survival as the old niches persisted, but as an opportunity for oxygen adapted organisms to now dominate. Aerobic prokaryotes quickly developed to occupy these new niches, and to exploit oxygen, and combined with cyanobacteria a new ecological system based around the utilization of oxygen developed.

Amongst these organisms was the ancestor of the mitochondrion. It is now widely accepted that this ancestor was most likely an α -proteobacteria, a free-living organism capable of independent metabolism using oxygen as an electron chain terminator (Emelyanov 2003; Thrash et al. 2011). This organism was endosymbiosed by the proto-eukaryote and transformed into the mitochondrion, however the stage at which this organism was adopted is still disputed in molecular biology, the two major theories of which are presented below.

1.2.2.1 Phagotrophy model

The 'phagotrophy model', primarily championed by Cavalier-Smith (Cavalier-Smith 2009), suggests that the acquisition of the mitochondrion was a late event in eukaryogenesis (Figure 1.2.2). In this model the trigger for eukaryogenesis was a transition to a predatory, phagotrophic lifestyle. It is speculated that the prokaryotic ancestor to the pre-eukaryote first underwent a loss of its outer membrane; this single membraned cell is described as 'neomuran'-literally new-walled, and then developed into the sister clades of archaea and pre-eukaryotes.

In Archaea the cells made modifications to the lipids within the remaining inner membrane, and often retained a cell wall, possibly in adaptation to thermophily. The pre-eukaryotes losing their outer membranes and rigid cell walls were able to embark on a new lifestyle. It was conjectured that if these cells were living in microbial communities, that there

would be some osmolar buffering of the environment (Cavalier-Smith 2009), which would mitigate the evolutionary cost of dispensing with their rigid cell walls, by the same token it was theorized that cells within these communities would be surrounded by potential 'prey', and the trophic benefits of exploiting members of the community drove pre-eukaryotes into a predatory lifestyle.

Divergence of 'Neomuran' from prokaryotes

- Loss of outer membrane
- Replacement of murein wall
- Divergence of archaea and eukaryotes

Development of eukaryotic ancestor

- Facultatively aerobic
- Large secretome
- Adaptation towards phagotrophy

PHAGOCYTOSIS

Development of eukaryotic features

- Endosomal system
- Cytoskeletal structure
- Development of the nucleus

ENDOSYMBIOSIS

- Enslavement of the α -proteobacteria
- Peroxisomal ATP/ADP carrier used to tap energy
- Host lost own oxidative capacity, through compartmentalisation of this function to the mitochondrion

ORGANELLOGENESIS

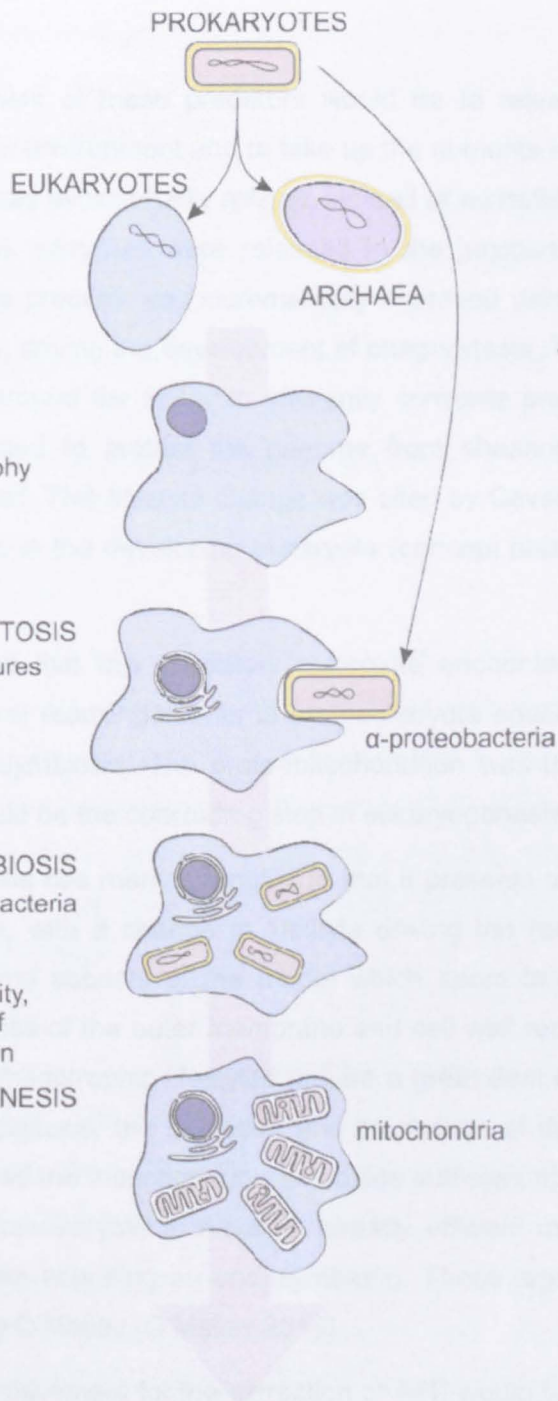


Figure 1.2.2 Phagotrophy model for eukaryogenesis

The phagotrophy model represents one current line of thought to explain the origins of eukaryotes, this figure illustrates the sequential events in the process, and is a simplification of recent work on this model by Cavalier-Smith (Cavalier-Smith 2009). The model theorizes that the eukaryotic ancestor was a sister clade to the Archaea, this organism, most likely living within a colony abandoned its cell wall and began a predatory lifestyle, acquiring eukaryotic characteristics to enhance its fitness as a phagotrophic predator. In this model the acquisition of the mitochondrion is a penultimate event in the process of eukaryogenesis.

The first development of these predators would be to release proteases and other digestive enzymes into their environment and to take up the nutrients released as they digested their prey. This approach was evolutionarily refined, instead of wastefully rejecting enzymes into the environment to diffuse, enzymes were released in the junctional contact with the prey organism. Efficiency of this process was incrementally improved with the ability of the cell to model itself around its prey, driving the development of phagocytosis. The endosomal system of these cells then evolved around the ability to efficiently consume prey and secrete enzymes, and the nucleus constructed to protect the genome from shearing as this predatory cell structurally re-arranged itself. This lifestyle change was cited by Cavalier-Smith to be crucial to driving the radical changes in the developing eukaryote (concept presented in (Cavalier-Smith 2009)).

It was at this point that this predatory eukaryote encountered the ancestor to the mitochondrion, the structural rearrangements to protoeukaryote enabled the cell to engulf the symbiont and begin endosymbiosis. The proto-mitochondrion was then exploited for energy production (ATP). This would be the concluding step in eukaryogenesis.

Whilst this hypothesis has many strengths in that it presents an orderly timeline for the evolution of the eukaryote, with a change in lifestyle driving the radical changes to cellular architecture, there are some aspects of the model which seem to run against evolutionary likelihood. Principally the loss of the outer membrane and cell wall represents an evolutionarily costly choice. In addition phagotrophic lifestyles require a great deal of energy, to dynamically reshape the cell for phagocytosis, the secretion and production of digestive enzymes. These activities might have required the mitochondrion to provide sufficient metabolic energy, or would have necessitated the protoeukaryote to have an already efficient metabolism (eliminating in part some of the gains from acquiring an endosymbiont). These arguments are presented in more detail in the review by O'Malley (O'Malley 2010).

Also the idea of enslavement for the extraction of ATP would have necessitated an ATP transport system to already have been present to offer an energy oriented advantage to maintaining the symbiont. It has been suggested that the mitochondrial endosymbiont possessed the gene (Emelyanov 2007) but prokaryotes with ATP exporters are rare, though some obligate parasites including the *Rickettsiales* have ADP/ATP transporters (Audia & Winkler 2006), whilst these normally work to extract ATP from the host, defective transporters might boost ATP in the host, and improve the host and its parasites survival. Alternately host

encoded transporters targeted through the endosymbiont outer membrane would either have to be innovated- perhaps from peroxisomal transporters, HGT from a *Rickettsia*-like parasite or from prior symbiotic events (O'Malley 2010).

The phagotrophy theory would also necessitate the existence of a population of amitochondriate pre-eukaryotes. Whilst many organisms have been cited as these eukaryotic progenitors (including *Trichomonas*) under the name of 'Archaezoa', none of these candidates has survived molecular biological evidence. Given the persistence of anoxic environments, as well as other primitive microbial communities elimination of the 'pre-mitochondriate' eukaryote through niche elimination does not seem likely (Embley & Martin 2006).

1.2.2.2 Metabolic syntrophies

The second major theory for eukaryogenesis (posed as the second major group in the review by O'Malley), posits that the acquisition of the mitochondrion was the key 'seed' event for the development of the eukaryote. Initially it was thought that a non-eukaryotic consortium of organisms, developed into an endosymbiosis (different models reviewed in (Embley & Martin 2006)). The exact members of this partnership are still subject to debate, and have created a myriad of syntrophy models, but the most persistently presented involve communities of archaea and prokaryotic eubacteria.

Hydrogen hypothesis

Of the earlier prominent theories, the hydrogen hypothesis, suggests that the original organisms were a hydrogen producing α -proteobacteria and an archaeal methanogen (Martin & Müller 1998). In this arrangement the hydrogen producing α -proteobacteria provided the hydrogen and carbon dioxide for the archaeal methanogen.

Methanogenesis is a form of metabolism which so far has only been observed in Archaea, which are thought to have emerged around 2.7 Gya (Hayes 1994) before the full extent of environment oxidation. These organisms are inhibited in the presence of oxygen as it competes with carbon dioxide as a terminal electron acceptor, thus this relationship must have originated in anoxic or microaerobic environments, or required the presence of a facultatively aerobic α -proteobacteria to mitigate the effects of environmental oxygen immediately around its methanogenic host. As the host became increasingly dependent on the endosymbiont for metabolism, firstly via endosymbiosis then as an organelle, it could then abandon methanogenesis, and pursue aerobic respiration- giving rise to the modern aerobic eukaryote.

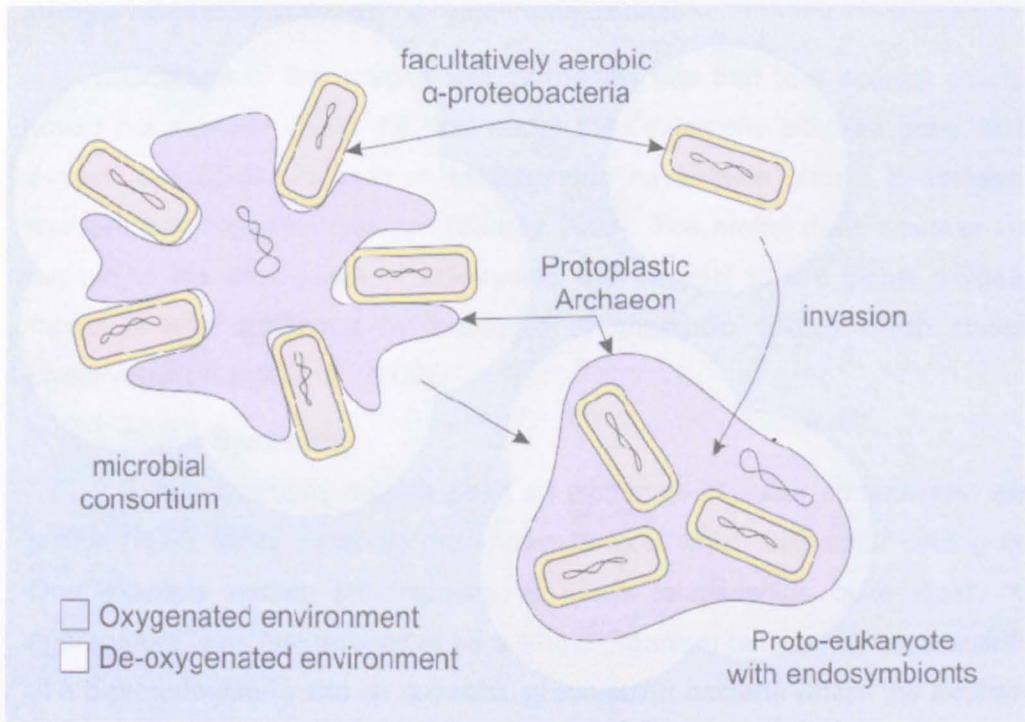


Figure 1.2.3 The oxygen toxicity model and microbial consortia

In several syntrophy models of eukaryogenesis the formation of an initial microbial consortium was stimulated from metabolic benefits, in the hydrogen hypothesis (Martin & Müller 1998), an archaeal methanogen would have benefited from a consortium with a facultatively aerobic hydrogen producing α -proteobacteria. In other models the α -proteobacteria would have consumed environmental oxygen enabling the archaea to maintain its metabolism in an increasingly oxic environment. This model has been adapted and built upon to include ROS mitigative mechanisms (Gross & Bhattacharya 2010).

This theory was energized by the discovery of hydrogen producing organelles-*hydrogenosomes*, in *Trichomonas* (Lindmark & Müller 1973), which were cited to fulfill the metabolic criteria suggested in the model. However the ancestral nature of the *Trichomonad* hydrogenosome is now disputed (Embley et al. 2003), as a likely secondary evolution of the mitochondrion, or the result of a different endosymbiotic event (Dyall 2004), whilst this erodes *Trichomonas* as an example of the primitive eukaryote, a hydrogen or anaerobic fermentative host oriented theory is still reasonable. In addition the methanogenic nature of the host has been criticized due to the lack of any trace of the original methanogenic metabolism, and no instance of any reversion to methanogenic lifestyle (Searcy 2003).

Oxygen Toxicity (OxTox) Model

The oxygen toxicity model initially introduced by Margulis (Margulis 1970), better developed further (Kurland & Andersson 2000) proposes that the oxygen mitigative effects of the symbiont would have been the 'service' which would benefit the host. In this model the

mitochondrial ancestor would have detoxified the host cytosol by converting oxygen to water using a respiratory chain with a cytochrome oxidase.

Criticism of the oxygen toxicity models cite that that internal aerobic endosymbionts would not logically shield the host better than externally adhered ones, or function better as oxygen detoxifiers, in addition mitochondria have been shown to release ROS potentially exacerbating oxygenic damage (Searcy 2003). The model does however have the benefit of explaining the emergence of eukaryotes with respect to the global oxygenation event. This model is also supported by examples of anaerobic ciliates which closely associate with prokaryotes (Yubuki et al. 2009).

Sulfur Syntrophy

Sulfur syntrophy models posit an exchange of sulfur compounds, especially hydrogen sulfide (H_2S). Many consortia are known to exist which use sulfur cycling in their symbioses. One example makes the intimacy of these relationships quite clear, “*Chlorochromatium aggregatum*” was first thought to be a single organism but was in fact a sulfur syntrophy pairing of a β -proteobacteria and phototrophic green sulfur bacteria where the exchange of sulfur is key to their symbiosis (Wanner et al. 2008). The role of sulfur is however criticized in eukaryogenesis because of the incompatibility of oxygen respiration with hydrogen sulfide, which would frustrate the path to extant aerobic eukaryotes from this consortium.

In general these different kinds of initial consortia would have immediately benefited, not from the exchange of ATP as in the phagotrophy model or in Margulis' original endosymbiotic model, but by the effects of metabolic efficiency, which gives a tangible evolutionary benefit for early endosymbiosis. To strengthen this argument, extant stable consortia of free living ciliates with prokaryotes have been discovered (as in *Calkinsea* (Yubuki et al. 2009)), replicating in part the kinds of relationships necessary to this eukaryogenesis model. However stability of this relationship would necessitate coordinated growth and expansion of the consortium.

These different arguments under the meta-idea of 'syntrophy' really reflect the uncertainty around the nature of the host (O'Malley 2010). Whilst each particular model of syntrophy reflects a different possible metabolic arrangement, each with individual strengths and weaknesses, it does present a fertile, testable ground to develop a eukaryogenesis model.

Criticism of the syntrophy models question broadly whether the original host was capable of engulfing the proto-mitochondrion. Whilst the phagotrophy model starts out with the development of a system to permit engulfment, the host within the syntrophy model is theorized

to be relatively primitive at the point it acquires the mitochondrion. However examples of prokaryotic cells with prokaryotic endosymbionts have been discovered. These include cyanobacteria with bacterial endosymbionts (Wujek 1979), and γ -proteobacterial endosymbionts, of the β -proteobacterial endosymbiont of mealyworm (von Dohlen et al. 2001). In unusual addition a genus of bacteria has been characterized with an intra-mitochondrial lifestyle (Sassera et al. 2006). These endosymbioses also fall against a backdrop of highly developed symbioses between prokaryotes in microbial consortia, for example "*Chlorochromatium aggregatum*" (Wanner et al. 2008).

Examination of archaeal vesicle fusion and budding systems provides support for phagocytosis as a functional possibility in early organisms. Homologues to the endosomal protein machinery have been detected in *Sulfolobus* (Ellen et al. 2009), as well as phylogenetic evidence showing monophyly between eukaryotes and a sub set of archaeal actin homologues (Yutin et al. 2009). These data could support an ancestral protoplasmic archaea (likely related to *Thermoplasma*) with actin cytoskeleton that could have engulfed the ancestral mitochondrial endosymbiont.

Alternately the host may have been 'invaded' by the mitochondrial endosymbiont instead (Dyall 2004). This relationship could have been commensal as previously stated or parasitic, the nature of which ameliorated when a drive towards metabolic efficiency allowed more cooperative relationships to prevail.

These arguments and evidence overcome some of the criticisms made by the phagotrophy supporters, and make the early acquisition of the mitochondrion a credible evolutionary path, possibly more conservative than the lengthy series of radical modifications proposed by the phagotrophy proponents. This archaea plus symbionts would then have undergone the redesign in cell architecture and the acquisition of other eukaryotic features, whose metabolic cost would have been compensated by metabolic enhancement by the proto-mitochondrial symbiont. This theory would also not necessitate a population of primitively amitochondriate proto-eukaryotes and is supported by examples of prokaryotic endosymbiontry.

The syntrophy model also attempts to explain the other eukaryogenic changes that emerged after endosymbiosis during the radical re-shaping of the archaeal host. For example genetic material transferred from the α -proteobacterial endosymbiont to the host archaea's genome has been suggested to instigate the formation of the nucleus. This theory cites that there was an emerging conflict as prokaryotic endosymbiont DNA was introduced into the host archaeal genome (Martin & Koonin 2006).

Suggested in Martin & Koonin 2006, eubacterial Group II introns from the symbiont would have been transferred and integrated into the archaeal host. As these introns replicated, evolved and dispersed within the host genome they would act to disrupt essential co-transcription/translation of host genes, or uselessly sequester translation/transcription machinery. In an attempt to segregate the spliceosomal activities required to remediate intron insertion, transcriptional and translational processes had to be segregated, giving rise to the nucleus as a transcriptional compartment, with translation occurring in the cytosol. However it has also been suggested that the archaea had already acquired eukaryotic characteristics by this time, in particular a sexual lifestyle to help mitigate the effect of intron spreading within the genome (Poole 2006).

Both models have to contend with issues about the chimeric nature of the eukaryotic genome, and variously propose different compositions of the pre-eukaryote, and the acquisition of genes from horizontal transfer from other organisms or the proto-mitochondrial endosymbiont. Additionally the hosts in both major models undergo loss of certain functions. In the phagotrophy model the host would have already been facultatively aerobic, and would have to lose this to the mitochondrion. In the syntrophy models the original host might have to lose characteristics such as archaeal membrane lipids or methanogenic metabolism.

The contestation of the events of eukaryogenesis is unlikely to be solved by any single piece of new data. However both models have to embrace the endosymbiosis of the mitochondrion ancestor. Whilst the order and impact of this event within eukaryogenesis is still being debated, the development of the mitochondrion had to necessarily confer benefits to the host.

1.2.3 Benefits of the mitochondrial endosymbiotic event

The two previously described models propose different routes for the first endosymbiosis to occur, but in both cases this was the first step in organellogenesis. This first endosymbiont then underwent the evolutionary transformation into the mitochondrion, and this section briefly considers which factors shaped the free-living α -proteobacteria into the mitochondrion, and why the mitochondrion developed into the type of organelle it did.

The first endosymbiont had to confer advantages to its host, advantages that would have to be maintained throughout its evolution into the mitochondrion. In the syntrophy model the ancestral endosymbiont would have conferred an immediate metabolic advantage to the host, for example mitigating environmental oxygen to permit methanogenic metabolism. The

ancestral α -proteobacteria is proposed to have been facultatively aerobic, and possessed cytochrome oxidases to employ oxygen as a terminal electron acceptor (Kurland & Andersson 2000). Whilst the host abandoned methanogenic or other fermentative metabolism, the mitochondria retained aerobic respiration, and continued to mitigate the effects of oxygen and reactive oxygen species. This model explains some key features of the mitochondrion, and the resulting metabolic organization of the eukaryotic cell.

The phagotrophy model posits the enslavement of the α -proteobacteria, and that the host evolved to extract ATP from the endosymbiont. ATP generation would directly benefit the host in a distinct way to the mitochondrion's effect at mitigating oxidative damage (Cavalier-Smith 2009). Whilst the capacity for the original proto-eukaryote to extract ATP is still debated, ATP production is an essential feature of the modern mitochondrion, and underpins the metabolic reliance of the host on its organelles.

The endosymbiont could have enhanced the metabolic efficiency of the developing eukaryote, even before the emplacement of nuclear encoded substrate transporters. By utilizing metabolic end-products of the host's metabolism the endosymbiont's other positive effects would have come at no cost to the host. Additionally metabolic processes in both systems could produce end products which are of value to either partner. Even though the metabolic systems of host and symbiont were not integrated, the energy available from a source molecule could be more efficiently utilised by the bipartite metabolism of the developing proto-eukaryote, than by either organism alone.

This improvement in efficiency, this would also apply in models whereby the endosymbiont was an invading 'parasite'. The combined organism would also be fitter in times of marginal energy abundance. Some reference to this effect has been cited in the bursts of evolutionary complexity around global glaciation events, whereby the difficulty faced by organisms placed a pressure on metabolic efficiency (Hoffman et al. 1998; Kirschvink et al. 2000).

Together these processes benefitted the host by providing aerobic tolerance, oxidative protection, metabolic efficiency and energy, and are features which have been conserved in the transition of endosymbiont to mitochondria. Whilst some benefits would have emerged early in this process the development of the mitochondrion into an energy producing organelle required the innovation of several new components.

For the mitochondrion to metabolically support the eukaryotic cell a system of transporters would need to be implemented in the membranes of the endosymbiont. The host would also need to regulate the endosymbiont's metabolic processes to benefit the whole cell, and would be achieved through adding regulatory factors to the metabolic machinery, particularly in the developing respirasome of the endosymbiont to regulate aerobic metabolism.

Endosymbiosis was simply the first step in organellogenesis, the endosymbiont had to confer a sustained benefit to the host throughout its transition into the mitochondrion. The eukaryogenesis models focus on specific aspects, such as oxygen mitigation, and ATP generation as indispensable features for the developing eukaryote, but other factors could have contributed to the maintenance of this relationship.

The persistent benefits of the ancestral endosymbionts meant retention within the evolving eukaryote. Living as an entity within the host eukaryote provided a special environment for evolution, and ultimately led to the unique modes of evolution that created the mitochondrion.

1.2.4 Endosymbiont to metabolic organelle

As soon as the endosymbiont began its symbiotic lifestyle a new set of factors contributed to its evolution, eventually culminating in its transformation into an organelle. The most immediate factor was that as soon as the endosymbiont began to propagate within its host and no longer participated in independent existence each eukaryotic cell became a 'genetic island'. During division only a few, perhaps even a single endosymbiont repopulated the daughter cells.

The evolutionary consequences of this population bottleneck are a classical genetic phenomenon, and lead to a progressive genetic deterioration, this effect is called 'Muller's ratchet' (Muller 1964). In this situation eukaryotic cells accumulate endosymbionts with deleterious mutations, a cell unlucky enough to receive defective endosymbionts has no way of replacing them in an asexual lifestyle. Cells with defective endosymbionts lose the advantages of their symbiosis, thus there was a selection pressure for the eukaryotic cell to maintain a functioning population of endosymbionts.

Concurrently there was a driving force for the eukaryotic cell to enhance the synergy of its biochemical interactions with its symbiont. Most obviously, and a key tenet in the phagotrophy model, is the leap made when the host can export ATP from its endosymbionts giving it access to new energy, and indirectly the wealth of the endosymbiont's metabolic

chemistry. These two forces to both maintain and exploit the endosymbiont's potential were defining factors in the development of the mitochondrion.

Supporters of the phagotrophy model have attempted to suggest that solute transporters enabling access to the endosymbiont's ATP could have been very quickly established through targeting pre-existing ATP transporters through the endosymbiont's endogenous outer membrane proteins (Cavalier-Smith 2006). These transporters would have emerged from a previous or concurrent endosymbioses or alternately from ATP transporters of peroxisomes. These transporters would then be used for the 'enslavement' of the mitochondrion (Cavalier-Smith 2006). However this model necessitates another population of extinct eukaryotes with these prior endosymbionts. Rather than invoke another population of ancestral organisms a simpler model whereby ATP was a significant benefit in a continuous timeline of endosymbiont/host syntrophy but not an initial or early factor is a more conservative model.

Whilst establishment of an inner membrane 'permeome' for the transport of energy and solutes might have been quickly achieved, the development of a protein targeting system penetrating to all endosymbiotic compartments would ultimately fulfill the host's need to regulate and maintain its endosymbionts and lead to the formation of an evolutionary state whereby symbiont degradation could be countered by nuclear transfer and host expression. A complete protein transport system, allowing informational flow from the host to the symbiont would require the innovation of protein transporters and complexes to cross the topological boundaries of the double membrane endosymbiont.

The foundations for this system would involve repurposing of existing proteins into the protein targeting system (Gross & Bhattacharya 2009), this would provide an evolutionarily expedient route for implementing the system, as well requiring the least evolutionary innovation. In addition to host proteins, a pool of proteins also existed within the endosymbiont. The ancestral symbiont necessarily had a complete protein export system, to target its endogenous proteins to its membrane compartments during its free-living existence. These proteins would be extensively re-purposed in the development of the import system.

Muller's ratchet describes a continuous subtractive evolutionary process which acts to degrade endosymbiont genomes (Muller 1964), but the eukaryotic cell was also shaped by other genetic processes. Genes from the endosymbiont were continuously available to the host (though a modified form of HGT, Endosymbiont Gene Transfer, EGT), through organelle rupture and autophagy, it can be assumed that although this material was integrated at a constant rate, there would be no selective pressure to conserve transferred material if it did not serve an

advantageous function. This additive process provided the eukaryote with a population of genes with which it could then develop and re-use. In addition the developing eukaryote could have received additional genetic material from Horizontal Gene Transfer (HGT) from independent organisms, or other transiently endosymbiosing entities. These sources could also be used to construct the protein import system.

The order in which the system was assembled is another point of contestation within the field of organellogenesis, and are fundamentally divided between two routes. First an inside-out(wards) approach, which has been partly detailed above in the proposal for early establishment of ATP transport (Cavalier-Smith 2006), whereby early establishment of ATP transport drove the development of a transporter import system, or an outside-in(wards) approach whereby the host assumes transport across the outer membrane, and then the inner membrane (Gross & Bhattacharya 2009).

1.2.5 Development of the mitochondrial preprotein translocases

The outside-in perspective (Gross & Bhattacharya 2009) sets a rational route for the establishment of host control, through small incremental benefits to the symbiotic partnership. Gram-negative bacteria face similar challenges to mitochondria and need to transport a variety of species across their membranes. In their evolution the double-walled gram-negative bacteria evolved a domain which was then employed to create carriers for their outer membranes. These carriers contain domains composed of repeated of anti-parallel β -strands wrapped to form a pore, this structure is called a β -barrel.

These proteins could then be adapted by changing the number of strands within a β -barrel by duplication of β hairpins and the insertion of residues into the pore conduit giving prokaryotes a flexible tool capable of transporting a variety of species from small molecules to polypeptides, either specifically or non-specifically (Koeblnik et al. 2000; Remmert et al. 2010). Amongst the families of β -barrel proteins developed in prokaryotes, β -barrel insertases were developed to insert these proteins into the outer membrane. As such the proto-mitochondrial endosymbiont already had a whole suite of specific pore forming proteins within its outer membrane, and the necessary machinery to deploy them.

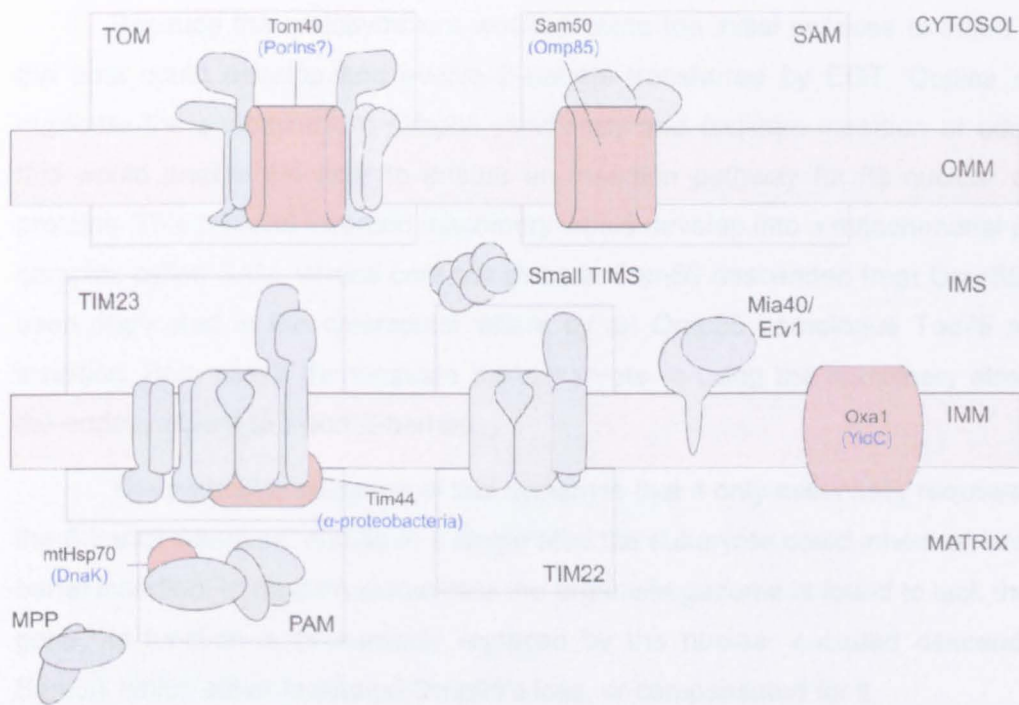


Figure 1.2.4 Endosymbiont contributions to the mitochondrial import model

The preprotein import system was synthesized from evolving proteins contributed by both the host (Green) and the Endosymbiont (Red) with the prokaryotic homologues indicated (Blue text). The outer membrane translocases have both been derived from prokaryotic β -barrel proteins (Gabriel et al. 2001; Paschen et al. 2003). Derived endosymbiont proteins are also found in the inner membrane insertase Oxa1 (Funes et al. 2011), and the Tim23 complex protein Tim44. The mtHsp70 is closely related to the prokaryotic DnaK, and has been shown to interact with Tim44 (Moro 2001) perhaps suggesting the conservation of a system which developed into the Tim23/PAM complex. Whilst the host innovated much of the translocase system, many of the essential components originate from the endosymbiont's existing machinery.

This machinery would be re-used in the eukaryotic cell. In the outside-in perspective (Gross & Bhattacharya 2009), establishment of eukaryotic control over the outer membrane was the first step in penetrating the membranes of the proto-mitochondrion, and the key proteins which the eukaryote was to re-use were members of the β -barrel family.

The Omp85 proteins are β -barrel insertases, these proteins mediate the insertion of β -barrels into the bacterial outer membrane, and they can facilitate their own insertion through previously inserted copies of Omp85 (Bredemeier et al. 2006). As such Omp85 molecules catalyze their own insertion.

As already discussed cellular events within the eukaryote/endosymbiont would continually provide the host with endosymbiont genes. Copies of endosymbiont β -barrels which successfully integrated into the nucleus would when expressed by the eukaryote generate proteins which would insert into the endosymbiont outer membrane through the existing endosymbiont Omp85 machinery.

Because the endosymbiont would provide the initial services to insert β -barrel proteins the host could develop and evolve β -barrels transferred by EGT. Copies of Omp85 would duplicate the endosymbiont Omp85 machinery and facilitate insertion of additional β -barrels, this would enable the host to ensure an insertion pathway for its nuclear encoded β -barrel proteins. This β -barrel insertion machinery would develop into a mitochondrial β -barrel insertase complex called SAM, whose core translocase Sam50 descended from Omp85. This process is seen duplicated in the chloroplast where by an Omp85 homologue Toc75 mediates β -barrel insertion. Both cases demonstrate the eukaryote re-using the machinery already employed in the endosymbiont to insert β -barrels.

The particular elegance of this system is that it only essentially requires a single protein, the β -barrel insertase, and so in a single step the eukaryote could inherit an entire system for β -barrel insertion. In modern eukaryotes the organelle genome is found to lack the original Omp85 gene, its function is presumably replaced by the nuclear encoded descendants Toc75 and Sam50, which either facilitated Omp85's loss, or compensated for it.

As mentioned in Gross and Bhattacharya's work (Gross & Bhattacharya 2009), the autocatalytic properties of the β -barrel insertase would quickly populate the endosymbiont membrane, and compensate for initial inefficiencies.

The conduit through which polypeptides could access the endosymbiont periplasm remain unclear. It is likely though that β -barrel proteins would have provided this conduit (Gabriel et al. 2001). In mitochondria β -barrel proteins were re-used to create a membrane spanning pore to facilitate the transfer of polypeptides into the periplasm. The β -barrel pore, a protein called Tom40 is theorized to have descended from prokaryotic β -barrel proteins, likely from a family of porins (Gabriel et al. 2001). The development of this protein was amongst the most significant steps in organellogenesis as it allowed the eukaryotic host an independent passage to transport proteins into the organelle.

Again autocatalysis is likely to have stimulated the establishment of eukaryotic Tom40 ancestral proteins, in conjunction with Omp85, or an ancestral eukaryotic Sam50, the insertion of Tom40 molecules would facilitate the translocation of more molecules of Tom40 from the cytosol. It has been theorized that the simplicity of these systems and their readiness to cascade autocatalyze led to the quick development of regulatory proteins in the SAM and TOM complexes (Gross & Bhattacharya 2009).

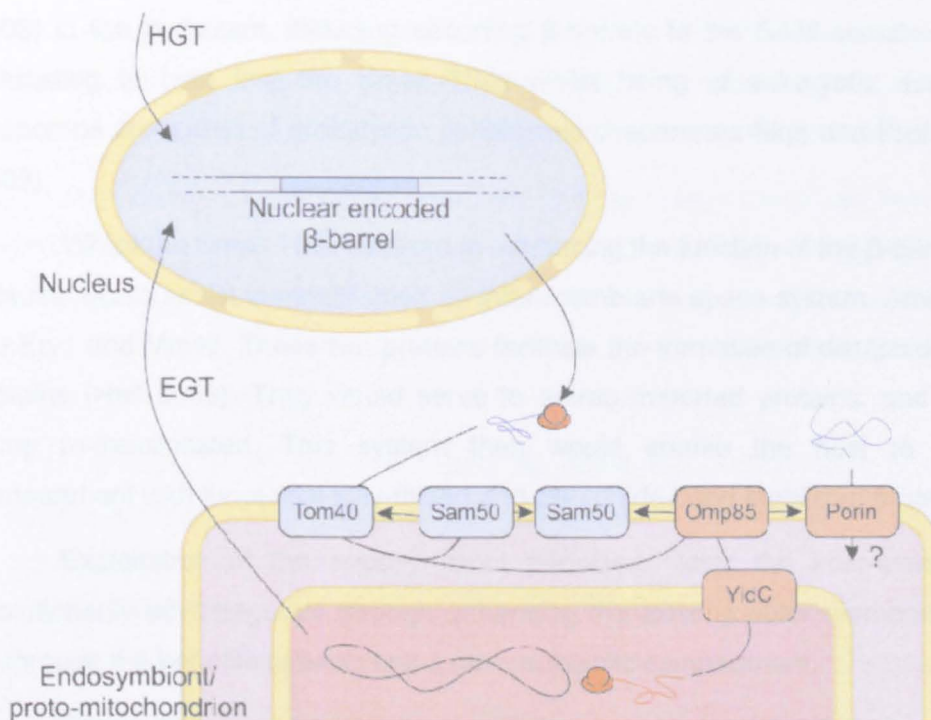


Figure 1.2.5 Establishment of host communication through the endosymbiont outer membrane

Nuclear acquisition of β -barrel genes from endosymbiont (EGT) or from other prokaryotic organisms (HGT) allowed the host to introduce new proteins into the outer membrane of the endosymbiont and establish communications through the outer membrane. The endosymbiont's pre-existing outer-membrane machinery would assist in establishing nuclear encoded β -barrels via insertion through Omp85. Copies of the Omp85 outer membrane β -barrel insertase would evolve into Sam50 which would functionally complement Omp85 and eventually replace it in the mitochondrion. Another β -barrel would develop into Tom40, an outer membrane translocase which conducts nuclear encoded polypeptides into the symbiont periplasm. If Tom40 had developed from prokaryotic porins (Gabriel et al. 2001), then endosymbiont porins might have provided the early access to the periplasm.

The establishment of eukaryotic control over the endosymbiont membrane would have brought evolutionary benefits, duplication of the β -barrel insertion pathway would have enabled the host to install new carrier proteins into the outer membrane, developed from symbiont carriers from EGT. These would enable the host to control the passage of small molecules across the outer membrane of the endosymbiont, perhaps enhancing the metabolic relation between host and symbiont. The β -barrel insertion machinery would also set the stage for a family of eukaryote innovated proteins to penetrate a new compartment of the endosymbiont-the periplasm (Gross & Bhattacharya 2009).

With the establishment of the TOM complex, the eukaryote could direct any protein to periplasm, regardless of its descent, prokaryotic or not. Some of the proteins innovated by the eukaryote and targeted to the periplasm helped again to establish and improve the β -barrel insertion machinery. Amongst these proteins are a series of small chaperones called the small TIMs. These proteins assist in transferring pre-proteins between translocases (Petrakis et al.

2009) in the periplasm, including escorting β -barrels to the SAM complex for insertion. It is interesting to note that the Small TIMs whilst being of eukaryotic descent duplicate the chaperone processes of prokaryotic periplasmic chaperones Skp, and Prefoldin (Petrakis et al. 2009).

Whilst the Small TIMs assisted in enhancing the function of the β -barrel machinery, other proteins would assist in establishing an inter-membrane space system. Amongst these proteins are Erv1 and Mia40. These two proteins facilitate the formation of disulphide bonds in imported proteins (Hell 2008). They would serve to entrap imported proteins, and prevent them from being re-translocated. This system then would enable the host to populate this new compartment with functional fully-folded and disulphide-bond stabilized proteins.

Exploitation of the endosymbiont periplasm, later the inter-membrane space, was evolutionarily advantageous through enhancing the existing outer membrane systems, as well as through the benefits of exploiting a new metabolic compartment.

Access to the inter-membrane space presented the inner membrane for the next step in the development of a membrane spanning translocase system. For theorists who propose an early establishment of eukaryotic proteins in the inner membrane (Cavalier-Smith 2006), eukaryotic proteins would have already accessed the inner membrane through navigating the outer membrane translocases, and encountering an inner membrane translocase.

In modern mitochondria the translocases responsible for the insertion of the mitochondrial ATP carriers do not appear to have arisen from prokaryotic proteins. The outer membrane demonstrates a clear example whereby the host re-uses the existing machinery to insert proteins of a similar type into the membrane. It might be expected then if an ATP transporter was inserted by a symbiont protein into the inner membrane then that endosymbiont insertase would be duplicated by the eukaryotic cell to maintain ATP carrier insertion (as in Omp85).

What is seen from the phylogenetic data is that the translocases on the inner membrane are much more of eukaryotic design (Kurland & Andersson 2000; Andersson et al. 2003). This development would suggest a sequential outside in perspective as suggested by Gross and Bhattacharya's work. As β -barrel proteins are ubiquitous to prokaryotes, the solute transporters of the mitochondrion are of eukaryotic descent, they feature repeated alpha-helices instead of β -strands, and are likely descendent of the eukaryotic endosomal system.

In order to populate the inner membrane with these host innovated proteins, new membrane insertases were also required. If the endosymbiont had the necessary insertase

which were compatible with eukaryotic transporters, then it would be expected that these insertases would be retained. Instead the translocase complexes of the inner membrane do not show proteobacterial descent, suggesting that the host innovated the translocases.

The likely ancestral translocase for the inner membrane would have been a Tim22 like protein. This protein in mitochondrial is a membrane insertase, and facilitates in the insertion of mitochondrial proteins into the inner membrane. Proposing a Tim22 like protein as the ancestral translocase of the endosymbiont outer membrane is also attractive as it can apply to the same principles of autocatalysis theorized for the outer membrane.

The Tim22 like ancestor would be capable of inserting copies of itself into the mitochondrial membrane, establishment of single molecules would then catalyze a self sustaining population of transporters within the inner membrane. This process would mirror the establishment of SAM in the outer membrane. A Tim22 insertase would then be responsible for populating the inner mitochondrial membrane with transporters, which arose through duplication and retargeting of existing eukaryotic proteins.

In as much as the TOM complex was a critical step in the establishment of a molecular dialogue between host and symbiont, the establishment of an inner membrane insertase was essential for the metabolic integration of the two organisms. The development of the eukaryotic 'permeome' provided the channel to exchange metabolites between cytosol and proto-mitochondrial matrix. The development of wholly eukaryotic transporters, and an independent path for protein insertion meant that the host could take control of the proteins it inserted into the inner membrane without reliance on a prokaryotic system. In this way the host could direct the metabolism of the proto-mitochondrion, and finally to extract ATP (Gross & Bhattacharya 2009).

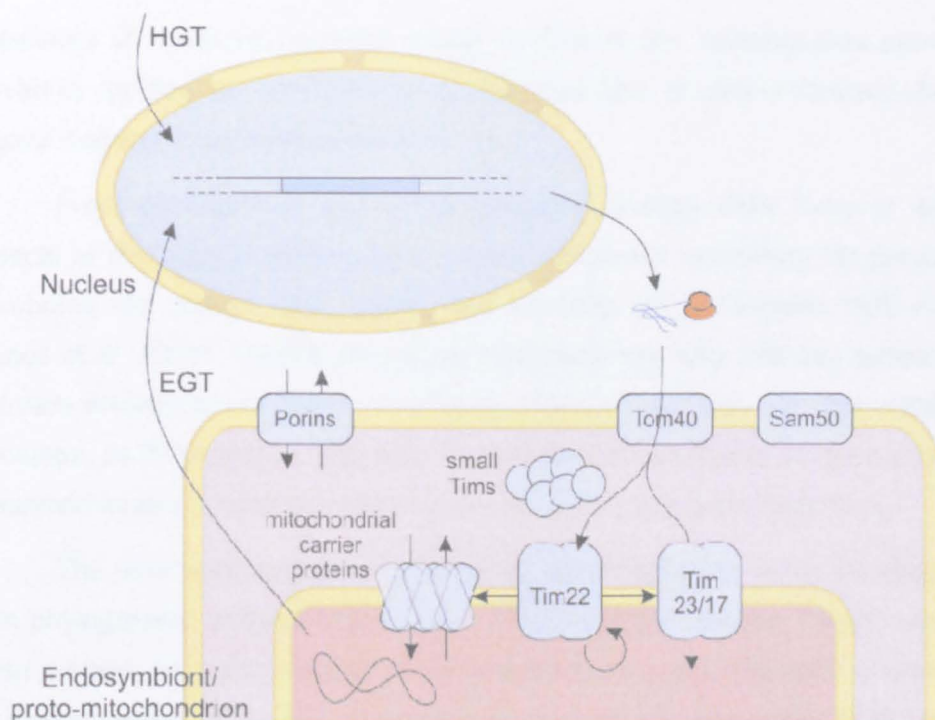


Figure 1.2.6 Establishment of permeome and inner membrane translocases in the proto-mitochondrial symbiont

Establishment of outer membrane translocases enabled the host to access the periplasmic space and symbiont inner membrane. Introduction of the small Tims, and Erv1/Mia40 complex aided in the insertion of pre-existing endosymbiont targeted proteins and provided a periplasmic chaperone pool. The Erv1/Mia40 complex would serve to entrap outer membrane translocated proteins within the IMS. Invention of an ancestral Tim22 like protein would establish an Inner membrane insertase, which would enable further ancestral Tim22 like proteins to be inserted as well as mitochondrial carrier proteins. Finally gene duplication and divergence of the Tim22 ancestral gene would have given rise to Tim23/17 proteins and translocase complexes able to import into the symbiont cytosol/proto-mitochondrial matrix.

Penetration to the final layer of the mitochondrion would mirror the establishment of the TOM complex in the outer membrane. As the β -barrel proteins accommodate a family of both pore-like, and insertase functions, the Tim22 ancestor could have acted as an insertase for pore like Tim23/17 proteins. Tim17/22/23 all share the same α -helical transmembrane domains, and have likely arose through gene duplication. Whilst Tim22 mirrors the function of Omp85 in the outer membrane, the functions of Tim23/17 mirror that of Tom40 in the TOM complex. These proteins would facilitate insertion of eukaryotic, and EGT derived genes into the mitochondrial matrix (Gross & Bhattacharya 2009).

The inner membrane translocases represent a system of eukaryotic innovation, but the endosymbiont must have had pre-existing machinery for inserting proteins into its inner membrane. As the host had to compensate for endosymbiont genomic degradation there must have been a necessity to replace and substitute the functions of this original system. The

installation of eukaryotic derived solute carriers in the mitochondrial permeome might have served to replace the endogenous symbiont carriers of similar function, but do relicts of the original endosymbiont system survive?

From phylogenetic and some molecular biology data there is some evidence that aspects of the original endosymbiont inner membrane machinery do persist. Within the inner membrane the protein Oxa1 shows descent from the prokaryotic YidC membrane insertase (Funes et al. 2011). Oxa1's involvement supports the idea that this protein was necessary to maintain endosymbiont inner membrane proteins, but was not the catalyst for the Tim22 revolution, as if this protein was able to catalyse the insertion of solute carriers the innovation of eukaryotic inner membrane translocases would not have been needed.

The other evidence for endosymbiont relict machinery within the inner membrane comes from phylogenetic analysis between the TIM23 complex protein Tim44, and mtHsp70. Both of these proteins have some evidence for prokaryotic descent, mtHsp70 is strongly similar to DnaK. In the mitochondrial system, Tim44 serves as a docking site on the TIM23 translocase to which a protein motor powered by mtHsp70 attaches. The interaction with Tim44 and DnaK has been probed experimentally (Moro 2001). This may then represent a prokaryotic system which has been repurposed to serve the TIM23 translocase.

1.2.6 Transition to mitochondrion and proteomic chimerism

The innovation of these membrane translocases were critical events which were essential to maintain the proto-mitochondrion against inevitable genomic degradation by Muller's ratchet, as well as metabolic exploitation by the eukaryotic host. The establishment of this system would also precipitate radical changes in the proto-mitochondrion. The emerging mitochondrion retained only a small fraction of its original genome, and was substantially reduced in its capacity outside of aerobic metabolism. However new host originated proteins would be introduced into the proto-mitochondrial proteome.

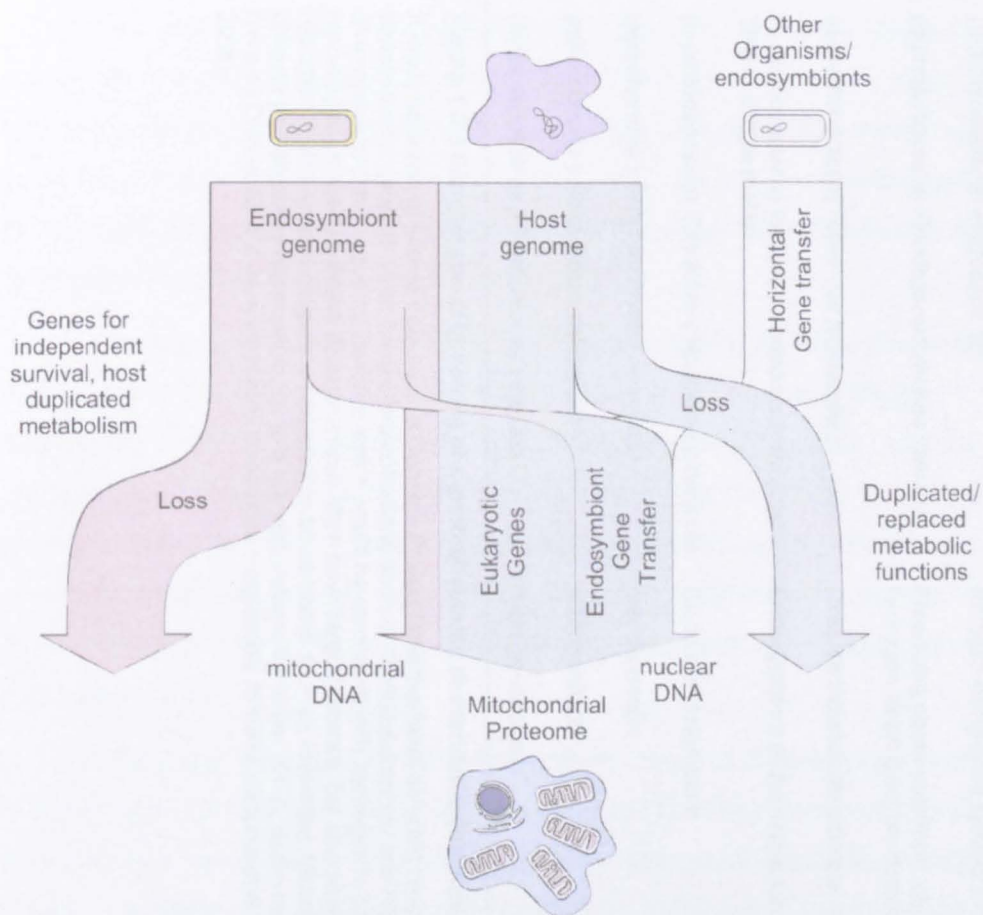


Figure 1.2.7 The evolutionary origins for the modern mitochondrial proteome

The host and endosymbiont originated with complete genomes, which have been significantly remodeled during the development of the mitochondrion. Whilst the genomic size of the original mitochondrion is unknown, free living α -proteobacteria have genomes of several thousands of genes, the mitochondrial relict genome encodes only tens to hundreds of genes (diversity of this is reviewed in part in (Adams & Palmer 2003)). Some of this genetic material was simply lost, such as genes encoding for independence, reliance on the host also gradually eroded away other functions in the symbiont. Some genes that were lost were substituted with host genes, and these were re-directed to the mitochondrion. With the establishment of the preprotein import complex endosymbiont genes transferred to the nucleus could be re-expressed in the mitochondrion. Additionally other proteins acquired from ancient horizontal transfer event could be expressed in the mitochondrion through the targeting system.

| Organism | Lifestyle | Genome size (Mbp) | Gene number | Proteome |
|--|--|-------------------|-------------|----------|
| (α -proteobacteria) <i>Methylobacterium nodulans</i> | Free-living aerobic, facultatively methylotrophic, nitrogen fixing bacteria | 8.8 | 8308 | - |
| (α -proteobacteria) <i>Rhizobium etli</i> | Nitrogen-fixing plant symbiont | 6.5 | 5963 | - |
| (α -proteobacteria) <i>Rhodopseudomonas palustris</i> | Free-living photo-autotroph, chemo-heterotroph, uses hydrogen as an electron donor | 5.7 | 5246 | - |
| (α -proteobacteria) <i>Rickettsia prowazekii</i> | Obligate intracellular parasite, multiple species | 1.1 | 833 | - |
| (α -proteobacteria) <i>Wolbachia</i> endosymbiont strain TRS of <i>Brugia malayi</i> | Endosymbiont of the nematode parasite, <i>Brugia malayi</i> | 1.08 | 1745 | 1610 |
| (α -proteobacteria) <i>Candidatus Hodgkinia cicadicola</i> | Cicadia endosymbiont | 0.14 | 169 | 169 |
| (mitochondria) <i>Saccharomyces cerevisiae</i> | mitochondrion | 0.08 | 46 | ~850** |
| (mitochondria) <i>Plasmodium falciparum</i> | mitochondrion | <6kb | ~3 | - |
| (hydrogenosomes) <i>Trichomonas vaginalis</i> | hydrogenosome | 0 | 0 | ~350 |

Figure 1.2.8 Comparison of genomes of α -proteobacteria, descended symbionts, and organelles

A small selection of entries from NCBI's Genome Project Data has been selected here for this table, showing the diversity in α -proteobacteria, and their descendent organelles. Free living α -proteobacteria such as *Methylobacterium* and *Rhodopseudomonas* exhibit varied metabolism showing autotrophy and chemoheterotrophy, including lifestyles which would be consistent with 'hydrogen hypothesis' type models, these organisms have large genomes. α -proteobacterial symbionts like *Rhizobium* can often have large genomes, but also small ones such as *Wolbachia*, and very small ones such as *Hodgkinia*. Genome reduction is also seen in obligate α -proteobacterial parasites such as *Rickettsia*. Mitochondria tend to have small compact genomes, ranging between the smaller obligate parasites and endosymbionts, to a few kbp (Adams & Palmer 2003). Some mitochondrially derived organelles such as the *Trichomonas* hydrogenosome have no DNA at all- but have estimated proteomes comparable to some mitochondria and symbionts (Carlton et al. 2007).*(Heazlewood 2004), **(Reinders et al. 2006),

The degeneration of the endosymbiont genome is not the same as decay. Functions necessary to maintain the evolutionary benefit of the symbiotic event must have been retained. Mitochondrial genes transferred to the nucleus are protected from further genomic degradation by Muller's Ratchet, and exposure to the metabolically energetic mitochondrial matrix (Embley et al. 1995). Not all genes were retained in the transition, those related to independent survival or host duplicated processes were lost.

The effect of genome reduction can be seen before the emergence of the mitochondrion, and is a general feature of endosymbionts and endoparasites (Figure 1.2.8) (Andersson & Andersson 1999). In obligate cellular parasites, such as *Rickettsia*, and *Mycoplasma* genome reduction has already shrunk their genomes to about one tenth of their free-living relatives (Figure 1.2.8). The same effects are observed in endosymbionts though with considerable range. Some endosymbionts are capable of limited independence and multiplication outside of the host (Tang et al. 1990) whilst others are obligate, with the mitochondrion being an ultimate extension of this.

Whilst the initial nature of the ancestral mitochondrial endosymbiont is often weighed between symbiosis and parasitism, the same evolutionary forces apply to both models, and genomic reduction would have begun before the establishment of preprotein translocase systems. A further reduction in genome would be observed in the transition from endosymbiont to organelle, reducing gene number by another order of magnitude, the reasons for this further loss are briefly detailed below but are discussed in considerable detail in review: (Adams & Palmer 2003)

Endosymbiont genes relating to independent living would have no loss penalty to the host, for example loss of the symbiont cell wall, and would conceivably enhance overall systemic efficiency against other competing eukaryotes, favoring a selection towards genomic streamlining (Andersson & Andersson 1999; Sakharkar 2004).

Further genomic losses, affecting metabolic pathways are compensated/tolerated by increasing dependence on the host to supply metabolic building blocks, or for processing waste products. Alternately genes lost from mitochondria may be substituted with host genes re-targeted to the organelle, or horizontally transferred copies of the mitochondrial gene.

Intracellular populations of mitochondria might also compete, with genomic reduction giving reduced organelles a replicative advantage (Andersson & Andersson 1999; Adams & Palmer 2003). Conversely beneficial mutations might propagate best in nuclear encoded genes.

A more efficient energy producing organelle might not necessarily have a competitive advantage over other intracellular organelles, and its benefits might be lost. However beneficial mutations within the nucleus have benefits to all organelles of its descendants, making it fitter to compete.. Over time the cumulative 'improvement' of mitochondrial genes might only be realized within the nucleus (Adams & Palmer 2003). These factors would contribute to the dramatic genomic reduction of the endosymbiont, to the vestigial genomes seen in mitochondria. Paradoxically there must be forces which prevent complete loss of organellar genomes, as the efficiency of eliminating transcription and translation from the mitochondrial compartment would conceivably de-convolute and streamline the efficiency of the organelle, some discussion of these factors are discussed in review (Adams & Palmer 2003).

Some indication to the factors of mitochondrial genome retention can be inferred from the constituents of the remaining DNA. In small genomes such as in *Plasmodium* the only surviving genes in the organellar DNA encode respiratory genes (Vaidya & Mather 2009).

From these data it can be inferred that these genes must remain in the mitochondrion, however the reasons for this are still conjecture (reviewed in Adams & Palmer 2003). Some cite that the regulation of these respiratory genes allows the host to very tightly control gene expression and thus finely regulate respiration (reviewed in Race et al. 1999)). Others cite that these proteins are exceptionally hydrophobic and may not be translocatable from the cytosol (Popot & de Vitry 1990). The constituents of these genomes will be re-examined later in divergent organelles.

The proteomes of modern mitochondria vastly exceed their genomically encoded proteins. The protein contents of the mitochondrion are then chimeras of both, mitochondrially encoded genes, nuclear transferred endosymbiont genes (EGT), as well as novel genes innovated, synthesized (from exon shuffling) or acquired (HGT) by the nucleus that are now directed to the mitochondrion. This effect gives rise to the mixed phylogeny of the mitochondrial proteome as well as the organization of its genes across two genomes.

1.2.7 Diversity of eukaryotic endosymbiotically derived organelles

Since the kind of ultrastructural techniques that made organellar biology visible, the researcher has also had to contend with extraordinary organellar diversity. If the proposed models of the last universal eukaryotic ancestor were mitochondriate, or proto-mitochondriate on the basis of molecular biology data, mechanisms are still needed to explain the variety of organelles present in extant eukaryotes,

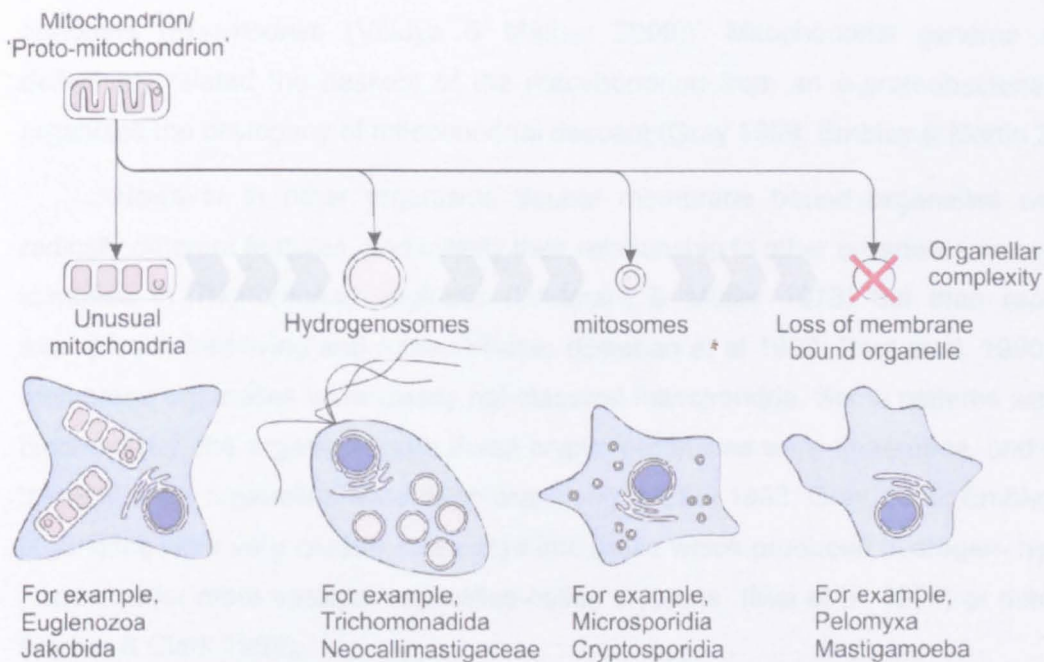


Figure 1.2.9 The diversity of mitochondrially derived organelles in extant eukaryotes

The proto-mitochondriate ancestor to all eukaryotes has since developed into a diverse group of eukaryotes with wide variety in their mitochondrially derived organelles. Often environmental factors such as anoxia have catalyzed the formation of highly divergent organelles from the proto-/mitochondrion, to better fulfill the metabolic needs of these organisms (Embley 2006). Some mitochondria exhibit only minor irregularities to their function or morphology, others are so radically departed from mitochondria that they have been given new names, hydrogenosomes and mitosomes.

From observation the majority of known eukaryotic taxa are mitochondriate, and have recognizable mitochondria with double membranes, mtDNA, oxidative phosphorylation and mitochondrial marker proteins. The consensual similarity of mitochondria in diverse taxa suggest a common origin in a mitochondrion like ancestor.

Some organisms have modified mitochondria- but are still recognizably mitochondrial, indeed within humans mitochondrial morphology can vary tissue by tissue (Calvo & Mootha 2010). Other organisms have structural deviations, such as discoid cristae observed in the Euglenozoa (Cavalier-Smith 1993). Others are remarkably reduced in metabolic capacity such the rudimentary respiratory chain in *Plasmodium* (Vaidya & Mather 2009), or have remarkable mtDNA, such as the multiple mini-circles of DNA present in kinetoplastids/trypanosomatids or the linear fragments in *Amoebidium* mitochondria (Marande et al. 2005; Marande & Burger 2007). However they are still recognizably mitochondrial.

Increasing amounts of molecular biology evidence, in particular mtDNA data revealed significant diversity within mitochondria. Some organisms retained relatively large mitochondrial genomes such as *Reclinomonas* (Lang et al. 1997) whilst others have atrophied relic mtDNA

genomes (*Plasmodium* (Vaidya & Mather 2009)). Mitochondrial genome data has also definitively related the descent of the mitochondrion from an α -proteobacteria ancestor, and organized the phylogeny of mitochondrial descent (Gray 1999; Embley & Martin 2006).

However in other organisms double membrane bound organelles were found with radically different features, and initially their relationship to other organelles was not known. First identified in *Trichomonas vaginalis* (Lindmark & Müller 1973) but then rapidly with other examples in free-living and rumen ciliates (Esteban et al 1993; Paul et al. 1990), these double membrane organelles were clearly not classical mitochondria. Some patterns were seen in their biochemistry, the organisms with these cryptic organelles were anaerobes, and many taxa with these strange organelles were deep branching (Muller 1993; Gray 2005; Embley 2006). These organelles were very crudely delineated into those which produced hydrogen- hydrogenosomes, and smaller more vestigial organelles called cryptons (Mai et al. 1999) or mitosomes (Tovar, Fischer & Clark 1999).

The presence of these alternate organelles promoted speculation on their origins (Dyall 2004; Embley 2006), which as mentioned above is still an active ground of contestation. With syntrophy models the last emerging eukaryotic ancestor was likely to be recognizably mitochondriate- if primitive. How then are these other organelles related, if at all to the mitochondrion. Are they products of the endosymbiosis of the mitochondrion, or in fact completely separate evolutionary events.

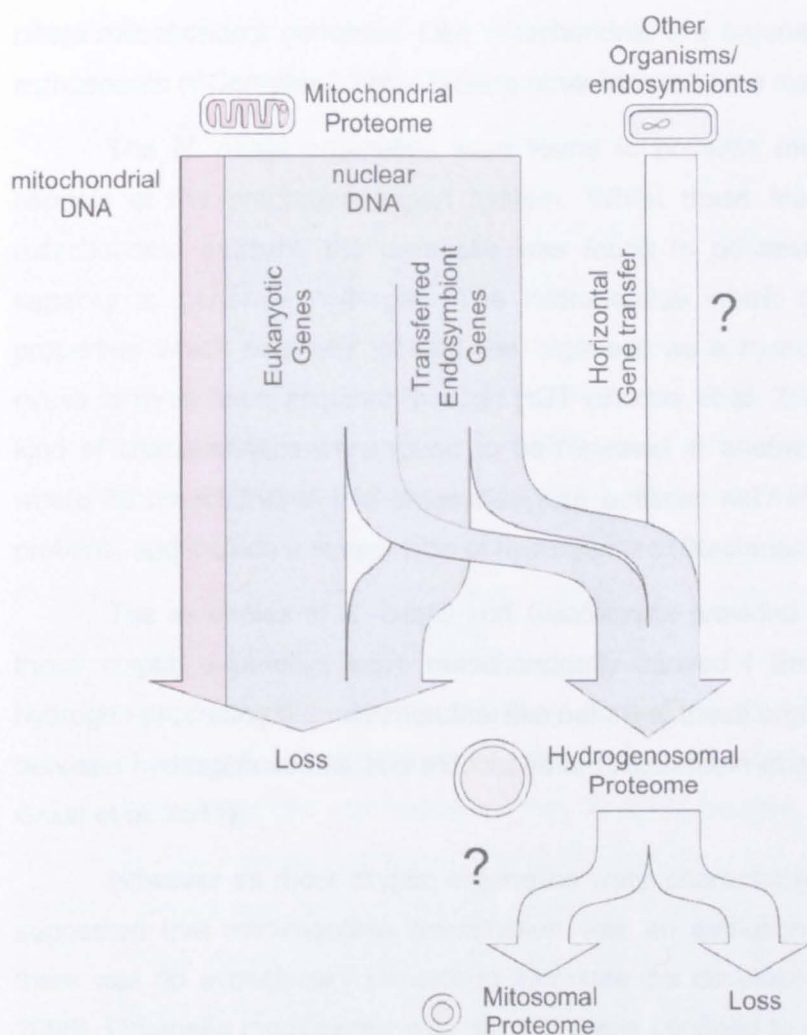


Figure 1.2.10 A model for the divergence of mitochondrially derived organelles

The divergent organelles observed in protists such as *T. vaginalis*, are widely theorized to have descended from mitochondria. The organelles are typically highly reduced, without genomes (Muller 1993). These organelles are also typically reduced in their metabolic capability (Muller 1988), though new genes may have been delivered to the evolving organelle from horizontal gene transfer in the eukaryotic nucleus (Boxma et al. 2007), though many of the unusual metabolic features of these organelles could have alternately been inherited from a more capable mitochondrial ancestor (Embley & Martin 2006). Even more reduced organelles such as the mitosomes of *Giardia* might have arose from continued degeneration of already divergent organelles, or have their own separate paths of evolution from mitochondria from other symbiotic events (Dyall 2004).

The cryptic organelles were generally beyond genetic analysis as most are without DNA, preventing the kinds of genetic analysis so successful with mitochondria. One exception to this rule was the characterization of the genomic DNA of a mitochondrial-like, but hydrogen producing organelle in the ciliate *Nyctotherus ovalis* (Akhmanova et al. 1998; de Graaf et al. 2011). The presence of genomic DNA within the organelle unambiguously confirms its descent from the mitochondrion. The *N. ovalis* organelle possesses a genome comparable to other

ciliate mitochondrial genomes. Like mitochondria, the organellar genome essentially encodes components of Complex I, but is lacking other factors of the respiratory chain.

The *N. ovalis* organelles were found to possess mitochondrial carrier proteins, and aspects of the preprotein import system. Whilst these features unambiguously indicate a mitochondrial descent, the organelle was found to possess new features, in particular its capacity to generate hydrogen. The hydrogenase which gave it the hydrogen producing properties which originally labeled the organelle as a hydrogenosome, was theorized in *N. ovalis* to have been acquired through HGT (Boxma et al. 2007; de Graaf et al. 2011). These kind of characteristics were found to be repeated in another unrelated species, *Blastocystis* where its mitochondrial like organelles also possess mtDNA, though with limited respiratory proteins, and include a similar type of hydrogenase (Stechmann et al. 2008).

The examples of *N. ovalis* and *Blastocystis* provided a growing body of evidence that these cryptic organelles were mitochondrially derived (Embley et al. 2003), and that the hydrogen producing but mitochondrial like nature of these organelles represented a missing link between hydrogenosomes and mitochondria (Stechmann et al 2008; Lantsman et al. 2008; de Graaf et al. 2011).

However as more cryptic organelles were characterised the polyphyly of their descent suggested that mitochondrial modification was an evolutionarily repeated process wherever there was an evolutionary process to stimulate the development of these organelles (Embley 2006). Organelle modification was not a process confined to evolutionary deep taxa either, but found in metazoans as in the facultatively anaerobic mitochondria of the *Ascaris* nematodes (Rew & Saz 1974), and *Paragonimus* flukes (Takamiya et al. 1994) and putative hydrogenosomes of the permanently anaerobic *Loricifera* of the meromictic hyper-saline L'Atalante basin (Danovaro et al. 2010).

In other examples the rapidity of organellar evolution was demonstrated when sister species within the genus were found to differ in organelle type, as in *Cyclidium* (Esteban et al 1993). These examples would suggest that hydrogenosomes can be rapidly developed when favorable evolutionary conditions are available (Embley & Martin 2006).

With an apparent multitude of polyphyletic divergent organelles to examine commonalities were sought to make sense of their descent and origins with respect to the mitochondrion. With strong evidence for an α -proteobacterial descent for the mitochondrion, genes from this family were sought in organisms with cryptic organelles. The results of these

studies were able to find mitochondrial proteins in these organisms, and frequently localized to the cryptic organelles- giving strong evidence for their relation to the mitochondrion (comprehensively reviewed in (Embley 2006). These data suggest that mitochondria, mitosomes and hydrogenosomes are commonly derived. Mitochondrial chaperone proteins Cpn60, Hsp70 are near universally discernible in these organisms, and prove to be an evolutionarily robust marker (Embley 2006).

Characterisation of specific cryptic organelles has revealed the presence of an ATP transporter of the MCF family, in hydrogenosomes-*Trichomonas*: (Tjaden et al. 2004) and in mitosomes- *Entamoeba*: (Chan et al. 2005) given the rarity of these kinds of proteins in the proteobacterial family, the presence of these proteins indicate that they likely descended from a complete mitochondrial system. The mitochondrial iron-sulfur machinery has also been identified in many cryptic organelles, *Trichomonas* hydrogenosomes (Sutak 2004), and in *Giardia* mitosomes (Tovar et al. 2003).

The composition of mitochondrial proteins within each cryptic organelle is however remarkably variable, and often difficult to discern given the speed of evolutionary change in some of the lineages. Hydrogenosomes and mitosomes represent proteomically reduced organelles with respect to the mitochondrion, and thus vary in the mitochondrial markers they harbor. Whilst these proteins indicate mitochondrial descent and represent a strong argument for the mitochondriate origins of these organelles. The presence of non-mitochondrial proteins and enzymes within these organelles has led to speculation of a mixed or hybrid origin (Dyall 2004).

Hydrogenosomes, are by definition organelles that produce hydrogen, though a few examples of hydrogen producing mitochondria have already been illustrated (*Blastocystis*). Metabolic energy in the form of energetic electrons which are ordinarily transported via carriers to the respiratory chain are instead diverted in hydrogenosomes to hydrogen producing enzymes called hydrogenases to maintain redox balance. This alternate route for regenerating the electron carriers obviates the need for aerobic respiration in these organisms.

Hydrogen production is often seen in anaerobic bacteria, but is an infrequent feature in eukaryotic mitochondriate aerobes, thus hydrogenosomes represent an uncommon but shared eukaryotic trait (Muller 1993; Embley 2006). Since the biochemical identification of the hydrogen metabolism of these organelles (Lindmark & Müller 1973), the proteins have subsequently been genetically characterised (*Trichomonas*; Bui & Johnson 1996, Horner et al. 2000, *Giardia* and *Entamoeba*; Nixon et al. 2003, *Nyctotherus*; Akhmanova et al. 1998, *Neocallimastix* and

Piromyces; Davidson et al. 2002, Voncken et al. 2002.) which allowed for phylogenetic analysis of the hydrogenosomal hydrogenases.

Investigation of the [Fe]-hydrogenases in these divergent organelles (Embley 2006) have indicated some monophyly between anaerobic eukaryotes with divergent organelles such as *Trichomonas*, *Entamoeba*, and *Giardia*, though this monophyly breaks down for other organisms such as *Nyctotherus*, and *Neocallimastix*. Whilst the exact relation of hydrogenosomal organisms like *Nyctotherus* with other eukaryotes cannot be fully determined from this one gene, the monophyly of other organisms suggest a common descent. All the hydrogenases detected in these organisms are most closely related to δ -proteobacteria, rather than to alternate eukaryotic or α -proteobacterial protein (Embley 2006).

Hydrogenases do not represent the only change to divergent organelle metabolism, another frequent substitution is switch from pyruvate dehydrogenase, to pyruvate: ferredoxin oxidoreductase (PFO) (Muller 1993). This substitution is seen in many of the amitochondriate organisms such as the *Trichomonas* hydrogenosome (Hrdý & Müller 1995), *Entamoeba* (Rodríguez et al. 1998), and *Giardia* (Emelyanov & Goldberg 2011). Again the identification of these genes from cloning or genomic data allowed for phylogenetic analysis of their descent. A similar trend of eukaryotic monophyly was observed, but the closest outgroups to the eukaryotic proteins are eubacterial, with no particular relation to α -proteobacterial genes (Horner et al. 1999). Additionally divergent organelles in the chytrid fungi group have evolved a separate method to circumvent PDH, and use a pyruvate: ferredoxin lyase (Akhmanova et al. 1999).

These data would suggest that some of the unique aspects of these organelles could possibly be vertically derived and were features of early mitochondriate organisms, though the origins for these genes seem unlikely to have been part of the original α -proteobacterial endosymbiont, but rather HGT events (these concepts reviewed with regard to PFO and hydrogenases in Embley 2006).

Whilst the proteomic chimerism of these organelles suggest a complex history of gene transfer from endosymbiont and other organisms, the organelles themselves are best explained by mitochondrial descent. The complexity of the mitochondrial import machinery, and other aspects of the mitochondrial metabolic biology (summarized and tabulated in Embley 2006) which have been shown to relate to these organelles are strong evidence that they are related to the original endosymbiotic event. Subsequent separate endosymbioses are not necessary to explain the chimerism of the divergent organelles, as originally amitochondriate eukaryotes are

unlikely to convergently evolve the fine detail of the mitochondrial system from a separate endosymbiosis.

Additional endosymbionts in already mitochondriate eukaryotes could have contributed to the proteome of the mitochondrial descended organelles. A mitochondriate host could conceivably re-use the mitochondrial import machinery to access secondary endosymbionts, and transform them into mitochondria like organelles (as duplicated import systems would target mitochondrial proteins to the secondary endosymbiont). However with all genetic data from these divergent organelles suggesting a relation to mitochondrial mtDNA rather than the genome of a separate prokaryote, a simpler model where genes are added to the mitochondrial proteome without the genesis of additional new organelles, or endosymbioses is simpler.

Transfer of new genes into anaerobic protist might have been facilitated by their lifestyle. Phagocytic protists adapting to anaerobic environments are likely to encounter, consume, and have access to the DNA of anaerobically adapted microbes, and rapid acquisition of environmentally relevant operative genes could have been possible.

Alternately a few intracellular obligate parasites are known in the lineages suggested by the PFO, and hydrogenase studies. For example *Lawsonia intracellularis* is a bacterial parasite related to δ -proteobacteria (and so *Desulfovibrio*) which infects eukaryotic cells. A similar parasitic agent could have been responsible for the δ -proteobacterial genes like the hydrogenases (Horner et al. 2000) being introduced into early mitochondriate eukaryotes.

The events that first created the mitochondrion in the eukaryote led to reshaping of the organelles proteome, and metabolic and genomic organization. These changes accommodated the evolutionary pressures at work on the endosymbiont and the host. The divergent organelles so far discovered reveal that the mitochondrion is still in the process of change, and that where environment dictates features can be lost, but also gained.

The hydrogenosome represents an evolutionary 'strategy' for anaerobic organisms. The monophyly of some enzymes (Embley 2006) suggest that these strategies might have been implemented in the ancestral mitochondrion to a wide group of organisms, or are the products of recent evolution.

1.3 *T. vaginalis* a protist with a hydrogenosome

This section takes a special look at *T. vaginalis* as a protist with a divergent organelle. The context of this organism with other 'deep-branching' eukaryotes is first discussed with

specific information on the commonalities of lifestyle and environment on its closest relatives (1.3.1). Next the clinical significance of *T. vaginalis* is discussed to highlight the urgency and potential benefits of research into this organism (1.3.2), which are developed to highlight the special role of the hydrogenosome (1.3.3). In this same section the existing knowledge collected about the hydrogenosome in *T. vaginalis* is displayed specifically addressing the remaining gaps in our knowledge about this organelle.

1.3.1 Evolutionary descent of *Trichomonas vaginalis*

This investigation will examine one of the better characterised protists with divergent organelles *Trichomonas vaginalis*. Firstly a brief mention of its evolutionary descent and lifestyles of its closest relatives will be presented.

Trichomonas belongs to a highly divergent eukaryotic clade near the eukaryotic crown (Sogin 1991; with more detail with respect to protists in Dacks & Doolittle 2001), and is grouped with morphologically similar flagellated protists in the Class Parabasalia. This class is proposed to belong to a suggested deep eukaryotic kingdom christened Excavata (Cavalier-Smith 2002). Whilst the monophyly of this group is uncertain, especially with regards to the Metamonads to which *Trichomonas* belongs (Hampl et al. 2009), many organisms classified as Excavate have unusual organelles.

Trichomonas then falls within a broad Phylum called the Metamonada, where it relates to other well studied 'amitochondriate' organisms such as *Entamoeba*, *Giardia*, and *Trimastix* which have all subsequently been attributed some mitochondrial descent (Reviewed and tabulated in Embley 2006, *Trimastix* by Hampl et al. 2008). *Entamoeba*, and *Giardia* possess highly reduced organelles called mitosomes, *Trimastix* has been found to encode genes for pyruvate ferredoxin oxidoreductase and hydrogenosome like hydrogenases amongst other mitochondrial proteins (Hampl et al. 2008). Within *Parabasalia*, *Trichomonads* branch from *Trichonympha* a genus of termite rumen protists with developed symbioses with spirochete type eubacteria (Gromov et al. 1977; Ikeda-Ohtsubo & Brune 2009).

Within *Trichomonads* a number of organisms have been characterized with similar gut lumenal lifestyles, including *Dientamoeba fragilis* (Windsor & Johnson 1999), *Histomonas meleagridis* (McDougald 2005), and *Mixotricha paradoxa* (Brugerolle 2004). Within the genus *Trichomonas* a number of species have been characterized again with gut lumen lifestyles such as *Trichomonas hominis*, but others have moved into new environments, *Trichomonas tenax*

(previously *buccalis*) is resident within the human buccal cavity (the mouth), whilst *Trichomonas vaginalis* has moved into the vaginal environment.

Evolutionarily *Trichomonas* belongs to a group of eukaryotes with unusual lifestyles. Many of its relatives have unusual organelles and thrive in anaerobic environments, often in symbiotic relationships with prokaryotes or within metazoan eukaryotes. Many of species of *Trichomonas* have parasitic lifestyles and cause pathogenesis in a wide variety of organisms throughout evolutionary history.

1.3.2 *Trichomonas* as a clinically significant pathogen

Whilst evolutionary and molecular biology aspects of the *Trichomonas* hydrogenosome form the primary subject of investigation in this thesis, some reference of course must be made to the clinical relevance of this organism in animal and human disease. *Trichomonas* appears to have been a parasite to animals throughout evolutionary history, it has been implicated in causing disease in distant extinct species (Wolff et al. 2009), and can be assumed to have at least colonised reptiles, avians and mammals ever since. As already stated, members of *Trichomonas* are causative agent of disease in many organisms. In animal husbandry species of *Trichomonas* are known to affect nearly all common livestock, *Trichomonas foetus/suis* in pigs and cattle, *T. ovis* in sheep, *T. equi / equibuccalis* in equines. The severity of colonization can be mild, or as yet unascertained, or have specific effects on reproductive ability such as in *T. foetus* (Felleisen 1999), or frank lethality, such as seen in avians infected with *T. gallinae* (McDougald 2005). As such the myriad species of *Trichomonas* represent a significant economic and welfare burden in the agricultural sector as well as ecological factor in wild populations.

At least three species of *Trichomonas* are known to colonise humans, *T. tenax* (previously *buccalis*) has an poorly characterised pathogenesis in the human oral cavity, though is associated with peri-odontal disease and respiratory tract infections . *T. hominis* which inhabits the gut lumen also has uncertain relationships to intestinal health (Chomicz et al. 2004). But the most clearly pathogenic is *Trichomonas vaginalis* which is the causative of Trichomoniasis.

Trichomonas vaginalis has developed a lifestyle likely descended from *T. hominis*. Like *T. hominis* it is actively phagotrophic (Garcia-Tamayo et al. 1978), and encodes a large secretome (Carlton et al 2007). Whilst *T. hominis* predaes on gut bacteria, *T. vaginalis* is somewhat larger enabling it to target human cells. Its cell remodelling processes give it the

capability to change shape to amoeboid forms to produce better contact with cell targets (Nielsen & Nielsen 1975). The downstream effects of *T. vaginalis* colonisation are to directly cause epithelial lesions through these processes. This results in inflammation, and provides infection paths for other opportunistic pathogens (Nielsen & Nielsen 1975; Garcia-Tamayo et al. 1978)

Whilst pathology is most apparent in females, the same pathological disturbances to epithelia can cause symptoms in males, resulting in subfertility, prostatitis and urethritis (Martínez-García et al. 1996). Whilst directly this epithelial erosion does not gravely affect health, its causal effects are significant. Co-infection of *Trichomonas vaginalis* is seen to facilitate transmission of other disease, and is a particularly important factor in the transmission of HIV in developing countries. There is some evidence that co-infection doubles the risk of HIV transmission, and given the large number of asymptomatic *T. vaginalis* infections, this pathway could represent an important pathway for disease propagation (Sorvillo et al. 2001). Co-infection is also correlated with adverse pregnancy outcomes and observed effects on infant mortality (Cotch et al. 1997). There is also some evidence that co-infection also increases the risk of cervical and prostatic cancer (Sutcliffe et al. 2009), possibly through the proliferative effects of the epithelia to counter lesioning, or providing an infection path for oncogenic papilloma viruses (Noël et al. 2010).

In summation *T. vaginalis* represents a causative agent of disease which remains widely overlooked in medical practice. Whilst affects of colonization are not so marked as those affected with other protest parasites, such as malaria, the incredible numbers of infected cases (reported >180million/year) make Trichomoniasis the most common non-viral sexually transmitted disease (Van der Pol 2007), as such it represents a large disease burden, especially in developing countries.

Trichomonas vaginalis' knock-on effects in HIV transmission and infant mortality also make Trichomoniasis an important factor in improving morbidity and life expectancy in developing countries, especially as the disease is easily remediated with proper medical attention.

1.3.3 The *Trichomonas* hydrogenosome

Like many other anaerobic eukaryotes *Trichomonas* has modified its metabolic organelles to adapt to anaerobic environments. *Trichomonas vaginalis* was the first identified with

hydrogenosomes (Lindmark and Müller 1973; Muller 1993), and work by these researchers set a foundation to which other divergent organelles would then be compared against.

The *Trichomonas* hydrogenosome has since been subject to extensive investigation partly from the primacy of the discovery of *Trichomonas* hydrogenosomes (Lindmark & Müller 1973), but also because the organism lends itself to culture and investigation (Diamond 1957). Electron microscopy work first revealed the membrane organisation of the organelle, and specifically identified the hydrogenosomal double membrane (Benchimol & De Souza 1983). This double membrane indicated that the hydrogenosome was not likely to be a novel endosomal system compartment, but perhaps of mitochondrial descent. This feature would then be revealed in other hydrogenosomes as a common feature (Muller 1993).

Biochemical analyses (Steinbüchel & Müller 1986) have also been employed to determine the metabolism of the hydrogenosome, and have revealed a system greatly divergent from typical mitochondria (Muller 1993). The principle metabolic functions of the organelle were first described by incubating isolated hydrogenosomes and were found to produce carbon dioxide and hydrogen, and exchange ADP for ATP (Steinbüchel & Müller 1986). The same experiments also began to define the metabolic pathways of the organelle, pyruvate was fermented into acetate, malate, and the gases hydrogen and carbon dioxide, in presence of ADP, phosphate, magnesium and succinate (Steinbüchel & Müller 1986). This provided a simple kind of 'black box' overview and indicated the hydrogenosome as an energy generating metabolic organelle.

The hydrogenosome in *T. vaginalis* was an important model in the investigation of mitochondrial derived organelles. The hydrogenosome exhibited properties consistent with some aspects of the mitochondrion such as ultrastructure and biogenesis, although sometimes employed different enzymes (for example PFOR in place of PDH), in other aspects the *T. vaginalis* hydrogenosome operated in a completely different way to the mitochondrion (for example hydrogen production).

As presented earlier in the introduction these essential differences in metabolism came to be seen as characteristic of these organelles, and provided common features into which organelles could be grouped.

The *T. vaginalis* hydrogenosome uses the alternate enzyme pyruvate: ferredoxin oxidoreductase (PFOR) instead of the mitochondrial pyruvate dehydrogenase, which uses ferredoxin to shuttle reaction electrons to an electron sink. Malate is also reversibly inter-

converted to pyruvate by a hydrogenosomal protein, malic enzyme (Drmotá et al. 1996), which also uses NADH to shuttle electrons from the reaction.

These unusual alternate enzymes are important because of their relation to the system through which the anti-protist drug metronidazole affects *Trichomonas* (Moreno & Docampo 1985; Quon et al. 1992). Electrons shuttled from these enzymes are conveyed by ferredoxin to a hydrogenase (Muller 1993), which recombines these metabolic electrons with protons to generate hydrogen. Metronidazole competes with the hydrogenase to accept electrons from ferredoxin, its action is seen in a corresponding drop in hydrogen production, Metronidazole is then activated into a radical form which then inflicts cellular damage on *Trichomonas* (Moreno & Docampo 1985).

The importance of understanding this mechanism is crucial to understanding the modes of resistance *T. vaginalis* is developing against metronidazole. One particular route for resistance arises through alteration in the expression of ferredoxin, or through the metabolic processes to which ferredoxin is coupled, principally PFOR (Cerkasovová et al. 1984; Kulda et al. 1993; D. M. Brown et al. 1999) The *T. vaginalis* hydrogenosome demonstrated reduction of the hydrogenosomal metabolism with respect to the mitochondrion. The tricarboxylic acid cycle from which most of the energy from pyruvate metabolism in aerobic organisms is extracted are fundamentally reduced to two enzymes in the *T. vaginalis* 'cycle', succinate thiokinase (STK) (Jenkins et al. 1991) and acetate: succinate CoA-transferase (ASCT) (van Grinsven et al. 2008). The interaction between these two enzymes cycle succinate and succinyl-CoA and produce acetate from the acetyl-CoA generated by PFOR.

The hydrogenosome has been found to lack the F₁/F₀-type ATPase (Lloyd et al. 1979) and so metabolic energy in the form of ATP has to be generated through substrate level phosphorylation, for example in STK.

Because of these novel metabolic pathways the *Trichomonas* hydrogenosome presents targets which are not found in the host, and so small molecule factors affecting these enzymes would demonstrate good specificity against *Trichomonas* without harm to the host.

However these alternate forms of metabolism share an organelle with pathways which are identical to mitochondria. Some essential processes are common to both organelles. The *Trichomonas* hydrogenosome has mitochondrial type iron sulfur machinery (Tachezy et al. 2001), divergent mitochondrial type ATP carriers (Tjaden et al. 2004), as well protein markers of some mitochondrial chaperones Hsp60,70 (Plümper et al. 2000).

These data suggest that the *T. vaginalis* has features of both mitochondrion and novel non-mitochondrial metabolism. However many aspects of the organelle remain enigmatic, and a collective theory to explain the origins of all the differences has not been arrived upon. To aid in this argument some fundamental properties of the organelle need to be discerned. These questions include the biogenesis of the organelle. Electron microscopy has already demonstrated that the hydrogenosome reproduces by fission of pre-existing organelles, and that like mitochondria the hydrogenosome has two membranes (Benchimol and De Souza 1983; Benchimol, Johnson, and de Souza 1996), however the proteins responsible for coordinating these processes remain unknown.

Also these organelles are known to be populated with a variety of proteins which have to be imported from the cytosol given the lack of organellar DNA. Whilst hints of the protein import system have been theorized from the chaperones Hsp60,70 (Plümper et al. 2000), the architecture of the pre-protein import system is also currently unknown.

New data from divergent organelles in other organisms such as *Entamoeba* have been characterized and show mitochondrial type machinery (Dolezal et al. 2010), though significantly reduced.

1.4 Why study the preprotein import machinery in the hydrogenosomes of *T. Vaginalis*?

The introduction has outlined the current research into the origins, evolution and divergence of mitochondrially derived organelles, and the mechanisms and forces that drove these changes to organelle and host.

This introduction has also put a special focus on membrane biology systems, and the necessity of them in membrane bound organelles, and the mechanistic aspects of their function. The introduction has also aimed to show the gaps in our knowledge about these subjects and unanswered questions to still be resolved.

To conclude this introduction, the objectives of the practical investigation will be presented to show why addressing the preprotein import machinery in the *T. vaginalis* hydrogenosome will contribute to all of these topics, and provide benefits from theoretical areas of contestation to practical biological data relevant to a poorly characterised pathogen.

This investigation will employ both computational informatics methods, to derive evolutionary data, but also importantly pursue the investigation in the physical world, by

characterising the molecular biology of a membrane translocase system, and looking at the composition of the hydrogenosomal membrane complex systems.

1.4.1 Why *T. vaginalis*?

T. vaginalis has been introduced as a disease causing organism, and a evolutionary missing link. This organism was selected in this investigation because it has unique properties.

From a practical perspective medical interest in this organism has fostered techniques to manipulate and culture this organism, and as such technical information related to its culture has been extensively characterised. Also medical isolates have been captured and deposited in biological repositories from which researchers have performed coordinated research on particular isolates. In addition an isolate of *T. vaginalis* have been sequenced (Carlton et al. 2007) providing the genetic information necessary to inspect the genome of *T. vaginalis* and to use molecular biology techniques targeting specifically encoded proteins.

As such *T. vaginalis* is a highly characterised organism, which can be cultivated and examined, with the benefit of prior research to refer to. Additionally, it serves as a useful model organism from which other researchers can continue research into endosymbiotically derived organelles. Aside from practical considerations the clinical importance of this prevalent pathogen, and its frequent neglect in research makes research in this organism worthwhile as the products of this research might contribute towards novel therapeutic strategies in the treatment of this organism.

Whilst the deep-branching nature of *T. vaginalis* is debated (Germot & Philippe 1999; Dacks & Doolittle 2001; Dacks et al. 2002) due to the relative speeds of evolution in disparate taxa do not necessarily enable the *T. vaginalis* hydrogenosome to be used as a tool to look into the past of organellar descent. *T. vaginalis* does represent a highly diverged eukaryote. Its relation to other mitochondriate eukaryotes can be examined to see the kinds of change that have happened, if not specifically into the timeline in which they happened. As such whether *T. vaginalis* is a deep branching eukaryote, or a recently rapidly diverging one, the processes shaping the changes to the hydrogenosome are still part of the process that changes one organelle into another.

1.4.2 Why the hydrogenosomes of *T. vaginalis*?

With emerging data on membrane translocases in other divergent organelles, the translocase in the definitive original hydrogenosome remain uncharacterised. Given that the *T.*

vaginalis hydrogenosome remains one of the more exhaustively studied divergent organelles, discovery of the divergent translocase system will fulfil this missing aspect of its membrane biology, and add to the information now available for other divergent organelles.

In addressing this question in the hydrogenosome, a proposed intermediate between the mitochondrion and the highly reduced mitosome, the hydrogenosomes of *T. vaginalis* represent a snapshot in the evolution of an organelle. Investigations have already characterised organelles at either side of this evolutionarily process, with the relatively mildly divergent organelles of *N. ovalis* and *Blastocystis* (Stechmann et al. 2008; de Graaf 2011), to the highly diverged and reduced mitosomes of *Entamoeba*. Whilst the evolutionary position of the hydrogenosome with respect to other organisms relies on better characterisation of the descent and rate of divergence of this particular species, commonalities between these organelles would suggest a common process of modification if not a common descent.

Again, practically, the *T. vaginalis* hydrogenosome has already been subject to extensive research, and particularly methods pertaining to purification, and for enzymatic markers for this organelle have been well represented (Muller 1993). This previous work makes the hydrogenosome of *T. vaginalis* practically accessible and provides an existing body of research to compare to. Similarly, future research building upon the work presented here can use these same approaches to further these advances.

From a cell biology perspective the hydrogenosome is the energy producing organelle of *T. vaginalis*, and is critical for cell viability. Clinical studies have already shown the importance of drug action on *T. vaginalis* with respect to the hydrogenosome (Lloyd & Kristensen 1985). This investigation will examine the hydrogenosome in detail, and can be used to define new markers, and new structural insights in the hydrogenosome which might be developed for detect the effects on new drugs against the *T. vaginalis* hydrogenosome.

1.4.3 The importance of the preprotein import machinery

The preprotein import machinery represents an intricate physical system, whose complexity and conserved common features make it an attractive target to examine the hydrogenosome's biology and descent.

Because the complexity of the mitochondrial preprotein import system is unlikely to have been duplicated precisely in a separate series of evolutionary events, as such the complete system is a reliable indicator of mitochondrial descent. In this sense the investigation of a whole system is a more thorough test for the hydrogenosome than the investigation of a single gene.

whose origin might be frustrated by *T. vaginalis*' rapid divergence, or have arisen through change horizontal gene transfer, as seen with investigations into the metabolic polyphyly of functions of this organelle (especially with regard to hydrogenases).

In *T. vaginalis* sequence divergence will provide an obstacle to finding homologues. However the mitochondrial import system contains proteins with distinctive domains found in no other eukaryotic system, and these can be used as markers and proxies for system homologues. Additionally the descent of individual members of the preprotein import system might reveal new aspects to the evolution of the mitochondrial import system if *T. vaginalis* is a deep branching organism.

Exploring the pre-protein import system also allows for a characterisation of the hydrogenosome across all of its compartments, reaching beyond the characterisation of previous studies. Protein entities discovered in this investigation would provide compartmental markers for future co-localisation studies, or serve to proxy information about the hydrogenosomal state in drug treated conditions.

Further than simply characterising the different compartments, the multitude of proteins composing the preprotein import system represent a population of indispensable proteins, and a process essential to the maintenance and biogenesis of the hydrogenosome, and thus cell viability. Characterisation of wild-type proteins, and their specific interaction partners and properties will provide data on the normal function of these proteins, and refine methods for the detection and purification, which might be of use in therapeutic applications designed to influence these proteins.

If phylogenetic analysis supports a true relationship between 'deep branching eukaryotes; and their preprotein import systems, agents designed to target the divergent machinery of their organelles would represent a new class of drugs with broad anti-protist activity, avoiding issues with existing drugs such as metronidazole which have general antibiotic activity.

1.5 Objectives of this investigation

The previous research in *T. vaginalis* provides a rich resource to construct new investigations, and allows for an investigation where multiple approaches can be employed to cross examine aspects of the hydrogenosomal translocase system. This investigation will also refine new techniques to work with *T. vaginalis* and as such will provide a resource to future research.

This project aims to identify members of the preprotein import system of the *T. vaginalis* hydrogenosome. Translocase candidates will be detected by both bioinformatic and molecular-biology methods. These two approaches will be brought together to address the preprotein transport system of the hydrogenosome. In addition to the identification of candidate translocase proteins, characterisation of these proteins *in situ* within the membrane will provide information to the precise localization of these proteins in the environments of their native function. The molecular biology aspect of this investigation will go beyond identification of candidates and will explore the architecture of membrane complexes in the hydrogenosomes of *T. vaginalis*.

Whilst the identification and characterisation of proteins within the *T. vaginalis* preprotein import system remain the core objectives throughout the investigation, the variety of methods employed to examine these proteins will also provide information peripheral to this objective, and inform more generally about the hydrogenosome, and ultrastructure of *T. vaginalis*.

These objectives will be devolved into the practical methods, and analyses within an investigative framework described in Chapter 3: Approach to practical work

2 Methods and materials

2.1 Chemicals and reagents

A short list of the suppliers for the reagents in this study are give on p235. Specific probes, antibodies, and enzymes are mentioned elsewhere in the methods text, with respect to their use.

2.2 Strains

2.2.1 *T. vaginalis* strains

In this study two *T. vaginalis* strains are used. G3 is the genome sequenced strain and was kindly donated by Prof. Graham Coombs (Strathclyde Institute of Pharmacy and Biomedical Sciences, University of Strathclyde). ATCC30001 (C1) is an un-sequenced strain extracted from a non-pathogenic clinical isolate, and was acquired from the ATCC repository, this strain was predominantly used in the hydrogenosome work. Growth of the C1 strain was found to be comparable to G3, and this isolates non-pathogenicity made this strain desirable to work with.

2.2.2 Bacterial strains

E. coli strains DH5 α , BL21 (DE3), BL21 (DE3) pLysS, were used in this study. Strain DH5 α was used for the cloning and maintenance of plasmid vectors. This strain has mutations in the endA1 endonuclease and EcoK1 restriction system, as well mutation of recA recombinase, these features make this strain suitable for transformation with un-methylated DNA, and maintenance of plasmids for which it was employed in this investigation. BL21 (DE3)/ BL21 (DE3) pLysS were used for induction of protein expression, these strains carry a prophage which expresses the T7 RNA polymerase controlled by a lacI system. These features allow for selective expression of plasmid inserts on IPTG induction

2.3 Vectors

2.3.1 *T. vaginalis* expression vectors

The plasmid encoding the cDNA for frataxin, TagVag, developed and generously gifted by P Dolezal (Department of Parasitology, Faculty of Science, Charles University in Prague), was used to express HA-tagged frataxin in *T.vaginalis* transformants. This plasmid was also the parent used to create other vectors containing HA-tagged proteins. Constructs made in these

vectors are able to be cultured in both *E. coli* and *T. vaginalis*, facilitating expansion of plasmid for final transformation in *T. vaginalis*. Selection is maintained in *E. coli* using ampicillin, and in *T. vaginalis*; G418. The essential features of these vectors are shown below.

| Name | Inserted <i>T. vaginalis</i> gene | Position of HA tag |
|--------|--|--------------------|
| pTV2 | Frataxin | C-terminal |
| pHBTv | Hmp43, <i>T. vaginalis</i> putative Sam50 | N-terminal |
| pHUP3A | Hup3a | N-terminal |
| pHUP3B | Hup3b | N-terminal |
| pHUP3C | Hup3c | N-terminal |
| pHUP3D | Hup3d | N-terminal |
| pHUP3E | Hup3e | N-terminal |
| pTIM44 | <i>T. vaginalis</i> putative Tim44 homologue | C-terminal |

Table 2.1 HA tag position within *T. vaginalis* expression vectors

Vectors based on the TagVag vector were developed in this investigation to express translocase candidates in *T. vaginalis* C1. Frataxin and Tim44 homologues are expressed as in TagVag with a C-terminal HA tag, for the β -barrel proteins this tag has been moved N-terminal to avoid disruption to the C-terminal β -barrel motif, which might effect membrane insertion.

2.3.2 Bacterial expression vectors

The parent vector for the expression of precursor proteins for *T. vaginalis* import assays was the pET26b vector (Novagen) and was used for subcloning *T. vaginalis* precursors in frame with a 6x His-tag. Selection was by virtue of the kanamycin resistance gene

2.4 Primers

| Name | Sequence | Construct |
|-------------------------|---|-----------|
| Precursor expression | | |
| AKFOR | 5'GGCCTTCATATGCTCAGTACATTAGCT-3' | pET26-AK |
| AKREV | 5'TTGGCCGGATCCCTGGAGGGCAGCATGAAGCTTC-3' | |
| FDXFOR | 5'GGCCTTCATATGCTCTCTCAAGTTTGCCGC-3' | pET26-FDX |
| FDXREV | 5'TTGGCCGGATCCGAGCTCGAAAACAGCACCATCG-3' | |
| Transformant constructs | | |
| HUP3AFOR | 5'-GCTTCTAGAAAATCTGGAAGTGAATATTATG-3' | pHUP3A |
| HUP3AREV | 5'-GCTCTCGAGTTAATTAGTGGAGTTTTCAACCC-3' | |
| HUP3BFOR | 5'-GCTTCTAGAATTTTGATCCAAAACAAAC-3' | pHUP3B |
| HUP3BREV | 5'-GCTCTCGAGTTATCGGAAGTTGTATCGC-3' | |
| HUP3CFOR | 5'-GCTTCTAGATTGAGTTCGTTCTTACATAC-3' | pHUP3C |
| HUP3CREV | 5'-GCTCTCGAGTTATTGGGAGTAAATCCTAGAC-3' | |
| HUP3DFOR | 5'-GCTTCTAGAATTCTTGAATCTTTTTCTCAAG-3' | pHUP3D |
| HUP3DREV | 5'-GCTCTCGAGTTATGTAGAAACAAAGTCATCTG-3' | |
| HUP3EFOR | 5'-GCTTCTAGATCCGAGCCTGAGCCTAAAC-3' | pHUP3E |
| HUP3EREV | 5'-GCTCTCGAGTTATTGGTTGATAGAAAGATCAATTC-3' | |
| TIM44FOR | 5'-GCTCATATGATTTGACGCTTTCTCAG-3' | pTIM44 |
| TIM44REV | 5'-GCTCTCGAGGGATCCAATTCTTGTTACTTGTGCTATCAG-3' | |
| HMPFOR | 5'-GACTCCCATATGTCATCTGCACCAGAGTGGTTC-3' | pHMP43 |
| HMPREV | 5'-GAGTCCCTCGAGAGCAGGAGTAATCCGAGCTGATA-3' | |
| Sequencing primers | | |
| pTagVag2For | 5'-GACATCGGCCACTTACGC-3' | |
| pTagVag2Rev | 5'-GTCAGCTGATGGTTCGAA-3' | |

Table 2.2 PCR primers used for vector production and sequencing

This investigation necessitates the creation of constructs for both the expression of His-tagged protein in *E. coli* as well as HA tagged protein to create *T. vaginalis* transformants. Two initial vectors were used to create the constructs in this study the pET26b vector (Novagen) was used modified to express *T. vaginalis* ferredoxin and adenylate kinase for exogenous expression, whilst the pTagVag2 vector (generously provided by P. Dolezal- see text) was used as the foundation to create the *T. vaginalis* transformation vectors. The primers used for the genetic manipulations in this study are shown above, as well as the primers used to sequence the finished constructs.

2.5 Cell culture

2.5.1 Bacterial cell culture

Cells were grown in Luria Bertani medium (LB), or LB supplemented with kanamycin (30 µg/ml) or ampicillin (100 µg/ml) as appropriate. Plates were prepared by the addition of 1% w/v powdered agar to the described media. All liquid cultures were grown with orbital shaking (> 200 rpm) at 37°C. Agar plates with bacterial colonies were stored at 4°C for up to 1 month.

2.5.2 *T. vaginalis* culture and maintenance

Cultures of *T. vaginalis* strain ATCC30001 (here in referred to C1) and G3 were maintained in supplemented Diamond's medium (Diamond 1957) (referred to here as TYM, 2% (w/v) tryptone, 1% (w/v) yeast extract, 14mM maltose, 11 mM L-cysteine, 1mM ascorbic acid, 6mM K₂HPO₄, 4mM KH₂PO₄) with 10% (v/v) heat inactivated horse serum, penicillin (100 units/ml) and streptomycin (100mg/ml), 179 µM ammonium iron sulphate, and 28 µM sulfosalicylic acid. Cultures were grown without orbital shaking at 37°C and typically passaged 1:10 into warm media every 1-2 days, when a cell density of >2x10⁶/ml was reached. Transformants were maintained in media supplemented with 0.5 µg/ml G418.

2.6 DNA manipulation techniques

2.6.1 Preparation of DNA

2.6.1.1 Genomic DNA preparation from *T. vaginalis*

Genomic DNA was extracted from late log cultures of *T. vaginalis* G3 of approximately 2.0 x 10⁸ cells (100ml). Parasites were pelleted by centrifugation (1500 x g, 10 minutes, 4°C) before being resuspended in lysis buffer (8 M urea, 2% (w/v) N-lauroylsarcosine, 150 mM NaCl, 1mM EDTA, 100 mM Tris-HCl, at pH 7.5). Nucleic acids were extracted using phenol: chloroform: iso-amyl alcohol, then precipitated in 0.6 volumes of isopropanol. DNA pelleted by centrifugation was washed in 70% ethanol, air dried, and resuspended in 10 mM Tris-HCl, 1 mM EDTA (TE buffer) and 10µg/ml RNAase at pH 8.0.

2.6.1.2 Plasmid preparation from bacterial cells by alkali lysis

Plasmid DNA was isolated from bacterial cells (DH5 α) using either commercial kits (Qiagen, Strataprep) or using an in house alkali-lysis procedure (modeled on the method developed by Birnboim & Doly 1979). In the latter, individual colonies were picked and transferred into 3ml of selective LB before overnight rotation at 37°C. Cells from 1.5ml of culture were pelleted by centrifugation (10,000 x g, 1 minute, RT) and resuspended in 200 μ l 50mM glucose, 10mM EDTA, 25mM Tris-HCl at pH 8.0 with 10 μ g/ml RNase A, cells were lysed by the addition of 200 μ l 0.2M NaOH, 1% (w/v) SDS and mixed quickly. Solution was then neutralized by the addition of 250 μ l potassium buffered 5M acetate solution at pH 4.8 which causes the precipitation of genomic DNA and SDS. The suspension is then cleared by centrifugation (16,000 x g, 10 minutes, RT), and the supernatant transferred and mixed with an equal volume of isopropanol to precipitate plasmid DNA. Purified DNA is then collected by centrifugation (16,000 x g, 10 minutes, RT) and washed with 70% ethanol before pelleting again under the same conditions. Plasmid DNA was then dried and resuspended in TE buffer.

2.6.2 Agarose gel separation of DNA

DNA samples were prepared in gel loading buffer (7% (w/v) sucrose, 0.04% (v/v) bromophenol blue, 10 mM EDTA, 0.02% (w/v) SDS, 3 mM Tris at pH 8.0) and analysed on agarose gels (1% w/v agarose in TAE buffer (40mM Tris base, 0.11% (v/v) glacial acetic acid, 0.04 % Na₂EDTA.2H₂O at pH 8.3) with 0.5 μ g/ml of ethidium bromide). Following electrophoretic separation DNA fragments were visualized on a UV transilluminator.

2.6.3 Phenol chloroform extraction

Extraction of purified DNA from mixed solutions was achieved by mixing aqueous solutions with equal volumes of phenol: chloroform: isoamyl alcohol (25:24:1) followed by vortexing. This process was repeated until no further impurities were precipitated at the aqueous interface. Trace quantities of phenol were removed from the aqueous solution by mixing with an equal quantity of chloroform. DNA in the aqueous solutions was then ethanol precipitated and re-dissolved in TE buffer.

2.6.4 DNA precipitation

DNA in aqueous solutions was precipitated by the addition 0.1 volume of 3 M NaOAc (pH 5.2), and the addition of 2 volumes cold ethanol. Solutions were then chilled at -20°C for at least one

hour. Precipitated DNA was then collected by centrifugation (16,000 x g, 10 minutes, 4°C), washed with 70% ethanol and pelleted again before resuspension in TE buffer.

2.6.5 PCR techniques

DNA sequences were amplified using Phusion DNA polymerase (NEB). A typical Phusion PCR reaction contains 1xPhusion HF Buffer, 1.25mM dNTPs, 0.5µM of each primer, and 2% (v/v) Phusion polymerase (1U/50 µl) with template DNA (typically ≥50ng). Typical cycling parameters were as follows; 30s at 98°C to melt templates and primer concatamers, followed by 25-30 cycles of 5s at 98°C, 30s at 50-65°C, 30s at 72°C. Final reaction intermediates were removed by incubating at 72°C for 5 minutes. Complete PCR reactions were then stored at 4°C

2.6.5.1 DNA sequencing using Big Dyes

DNA sequencing was performed using Big Dye™ v3.1 (Applied Biosystems). Sequencing reactions were set up to a final volume of 10 µl, containing 2 µl of 5x Sequencing buffer, 2 µl of BigDye™, primer to concentration of 1µM and 50ng of DNA template. The cycle conditions for the reaction are shown below. Sequencing reactions were subsequently analyzed by the Biopolymer Synthesis and Analysis Unit (BSAU, University of Nottingham).

| Program phase | No. of cycles | Temperature | Time |
|---------------|---------------|-------------|------|
| 1 | 1 | 96°C | 30s |
| 2 | 25 | 96°C | 30s |
| | | 50°C | 15s |
| | | 60°C | 4m |
| 3 | 1 | 4°C | Hold |

Table 2.3 BigDye sequencing reaction profile

The BigDye Reaction mixtures were used in conjunction with sequencing primers (see

Table 2.2) to verify construct sequence and fidelity. A typical reaction program is illustrated above, in the first step DNA and primers are denatured, before undergoing repeated sequencing cycles, finished reactions are then held at 4°C.

2.6.6 Restriction digestion

Restriction digests were performed as required with enzymes BamHI, NdeI, XbaI and XhoI (New England Biolabs, NEB) as per supplier's instructions. Reactions were assembled to 2-5U / μ g DNA and incubated at 37°C for 1-2 hours.

2.6.7 Ligation of DNA fragments

Restricted DNA fragments were ligated using T4 DNA ligase (NEB). Reactions were set up with a stoichiometry of 3:1 of insert to vector ratio containing ligase buffer (NEB) and ddH₂O to final volume of 10 μ l. Assembled reactions were allowed to run overnight at 16°C

2.6.8 Transformation of bacterial cells

2.6.8.1 Preparation of competent cells

1ml of strain overnight culture in LB broth was used to inoculate 100ml of LB broth (supplemented with antibiotics if appropriate), and cells were grown at 37°C with shaking until a OD₆₀₀ of 0.3-0.4 was reached. The cells were then collected by centrifugation at 4,000 x g for 10 minutes at 4°C, before being resuspended in 10ml of 0.1M MgCl₂ and incubated on ice for 2 hours after which cells were pelleted at 4,000 x g for 10 minutes at 4°C before being resuspended in 4ml 0.1M MgCl₂/10% glycerol and stored at -80°C.

2.6.8.2 Transformation of competent cells

Competent cells were incubated on ice for 20 minutes with vector DNA after which cells were heat-shocked for 60 seconds at 42°C before being replaced on ice for 5 minutes. Cells were transferred into 900 μ l of LB broth and incubated for 1 hour with shaking at 37°C before plating onto LB-agar supplemented with appropriate antibiotics.

2.6.9 Transformation of *T. vaginalis* cells

2.6.9.1 Preparation of *T. vaginalis* cells

T. vaginalis cells grown in TYM to $\geq 2 \times 10^6$ cells /ml and subsequently pelleted by centrifugation (1500 x g, 10 minutes, 4°C), and resuspended in fresh TYM to a density of 8×10^8 cells/ml. Aliquots of cells (300 μ l) were transferred into 0.4cm BioRad electroporation cuvettes.

2.6.9.2 Transformation with purified plasmid

10-50µg of DNA was mixed with the cells and left to incubate for 15 minutes on ice. The cell suspension was then electroporated using a Gene Pulser II set at 350V, 960µF. The cells were then resuspended in complete TYM at 37°C for 4 hours, after which G418 was added to a concentration of 50 µg/ml. After 21 hrs of incubation at 37°C the surviving suspended cells were transferred to a new conical tube, pelleted by centrifugation and resuspended in selective complete TYM.

2.6.9.3 Maintenance of transformants

Transformants typically took between 3-21 days to reach a density at which they could be passaged, and were passaged in the same manner as untransformed strains between 1:5-10 daily in G418 supplemented media.

2.7 Protein manipulation techniques

2.7.1 TCA protein precipitation

0.1 volumes of 100% tri-chloroacetic acid (TCA) was mixed with aqueous protein samples and heated for 5 minutes at 60°C, before being transferred onto ice for a further five minutes. Precipitated protein was pelleted by centrifugation at 16,000 x *g* for 15 minutes at 4°C. Pelleted protein was then washed with either ice cold acetone or ethanol, before recollection by centrifugation under the same conditions. The protein precipitate was air dried and subsequently resuspended in protein sample buffer.

2.7.2 SDS polyacrylamide gel electrophoresis

Protein samples were resuspended in sample buffer (10% glycerol, 50 mM TrisCl pH6.8, 2 mM EDTA, 2% SDS, 100 mM EDTA, 0.005% bromophenol blue), and loaded onto 10-15% polyacrylamide gels with a 4% stacking layer. The resolving gels were buffered in 375 mM TrisCl, pH 8.8, 0.1% SDS, and the stacking gel 125 mM TrisCl, pH 6.8, 0.05% SDS. Proteins were migrated into the stacking layer at a constant 100V before migration through the resolving layer at 150V.

2.7.2.1 SDS polyacrylamide gel staining

SDS PAGE gels were fixed by immersion in a 10% v/v methanol, 5% v/v acetic acid, 0.0001% (w/v) Coomassie brilliant blue solution until adequate staining was achieved, contrast was

enhanced by incubating gels in colourless destain solution (10% v/v methanol, 5% v/v acetic acid).

2.7.2.2 SDS polyacrylamide gel staining by Instant Blue™

Protein on polyacrylamide gels was visualized using Instant Blue™ (Expedeon), a colloidal Coomassie stain using the manufacturer's instructions.

2.7.3 Blue native polyacrylamide gel electrophoresis

Hydrogenosomal proteins suspended in solubilization buffer were mixed with 10x Blue Native (BN) PAGE buffer (5%(w/v), G-250 Coomassie Brilliant Blue, 0.75 M 6-aminocaproic acid, 100 mM BIS-TrisHCl, pH 7.0) to 1x concentration and loaded onto either 12% uniform BN PAGE gels, or 6-16% gradient gels, with a 4% stacking gel. Gel solutions were buffered with 50 mM BIS-TrisHCl, pH 7.0, 0.5 M 6-aminocaproic acid, and supplemented with glycerol in gradient gels to assist in proper mixing. The gel system anode buffer consisted of 50 mM BIS-TrisHCl, pH 7.0, and cathode buffer 50 mM Tricine, 15 mM BIS-TrisHCl, pH 7.0. The gel was initially run with the cathode buffer supplemented with 0.02%(w/v) G-250 Coomassie Brilliant Blue, which was then replaced with a colourless cathode buffer to better visualize migrating protein complexes.

2.7.4 Transfer of proteins to PVDF membranes

Polyacrylamide gels were blotted onto methanol wetted PVDF membranes in Geneflow semi dry blotter transfer system. Individual gels were stacked with PVDF and flanked by Whatman paper soaked in transfer buffer (25 mM Tris, 52 mM glycine, 20% (v/v) methanol). Transfer to the membrane was facilitated using a constant current of 480mA for 50 minutes.

2.7.5 Induction of protein expression in BL21

Overnight cultures (5 ml) were prepared from BL21(DE3) transformants using single colonies from freshly transformed plates, and 1ml was used to inoculate 100ml of antibiotic supplemented LB broth and grown with shaking at 37°C to an optical density of 0.5 - 1.0. At this protein expression was induced by adding IPTG (Isopropyl β -D-1-thiogalactopyranoside, Calbiochem) to a concentration of 0.5 mM. Cells were pelleted 3 hours after induction at 5,000 x g for 10 minutes at room temperature, and the pellets stored at -80°C.

2.7.6 Preparation of ^{35}S labeled protein precursors

Colonies of BL21 transformants were picked and grown overnight in a minimal growth media (M9; 2 $\mu\text{g/ml}$ Thiamine, 0.2% (w/v) glucose, 100 μM CaCl_2 , 9mM MgSO_4) supplemented with appropriate antibiotic. 0.5ml of culture was then used to inoculate a further 5ml of M9 selective media, and was allowed to grow with shaking at 37°C for around three hours until optical density reached 0.4-0.5. At this point protein expression was induced by the addition of IPTG to a final concentration of 0.5mM. In T7 polymerase containing expression systems rifampicin was added to a final concentration of 150 $\mu\text{g/ml}$, and 25 μl of ^{35}S labeled methionine was added (~44Tbq/mmol). The induced cells were incubated for three hours with shaking at 37°C, before being pelleted by centrifugation (5,000 x g 15 minutes, at 4°C). Pelleted cells were then processed by His-tag denaturing purification (2.8.9).

2.7.7 Detection of radio-labeled proteins in SDS polyacrylamide gels

Coomassie-stained PAGE gels were dried on top of Whatman paper under partial vacuum at 80°C for a period of >1hr. This process thins the gels and greatly increases the signal intensity of embedded labeled proteins. Dried gels were then laid against photographic film and exposed >24hrs in a sealed cassette.

2.7.8 Protein purification by His-tag under denaturing conditions

Pellets of induced cells were resuspended using a syringe at 10ml buffer / g wet pellet weight with 1%(v/v) Triton X-100, 6M guanidine hydrochloride, 0.1M NaH_2PO_4 , 0.01M Tris at pH8.0. The lysate was then rocked for 1hr at RT. Lysates were cleared by centrifugation (18,000 x g for 20 minutes at 4°C) and mixed with 100 μl of pre-equilibrated Ni-NTA His bind super-flow beads (Novagen) / 100ml of cell culture. Beads were then incubated for 90 minutes at room temperature with rocking, or over night at 4°C with rotation. Beads were recovered by centrifugation (1,000 x g, 5 minutes, RT). Non-specific protein was removed by a series of washes. The beads were first washed in 10 bead volumes of 1%(v/v) Triton X-100, 6M guanidine hydrochloride, 0.1M NaH_2PO_4 , 10mM Tris at pH8.0, then with 8M Urea, 0.1M NaH_2PO_4 , 10mM Tris, 20mM imidazole at pH8.0, and washed three times with 8M Urea, 0.1M NaH_2PO_4 , 10mM Tris, 20mM imidazole at pH6.3. Nickel bound protein was eluted by washing with 2 bead volumes of 8M Urea, 0.1M NaH_2PO_4 , 10mM Tris, 250mM imidazole at pH4.5, until all protein had been eluted.

2.8 Organelle preparation

The following techniques are also reproduced in a concise visual format on p95, and are in part developed from the organelle purification strategies employed by Bradley (Bradley 1997).

2.8.1 Preparation of crude hydrogenosomes and cytosol extract

Dense cultures of *T. vaginalis* strain C1, approximately 2.0×10^6 cells/ml, were pelleted by centrifugation ($1500 \times g$, 10 minutes, 4°C), and were then washed with 250mM sucrose, 20mM 1M HEPES (SH buffer) supplemented with 10mM β -mercaptoethanol, and separated by centrifugation under the same conditions. The pelleted cell mass was resuspended in 700 $\mu\text{l/g}$ cell pellet in SH supplemented with 5mM DTT, and protease inhibitors (25 $\mu\text{g/ml}$ TLCK and 10 $\mu\text{g/ml}$ leupeptin) (SHDI). The suspension was then passed repeatedly through a 23 gauge needle until $>90\%$ of cells were lysed, as evaluated by microscopy. Unbroken cells and cellular debris were separated by centrifugation ($1000 \times g$, 5 minutes, 4°C), and the supernatant was further centrifuged to produce a crude hydrogenosomal pellet (referred to here as the P1 hydrogenosomal pellet) by centrifugation ($8000 \times g$, 10 minutes, 4°C). The supernatant from this second centrifugation was further clarified by centrifugation ($16000 \times g$, 10 minutes, 4°C) to produce crude cytosol.

2.8.2 Preparation of purified hydrogenosomes

2 litres of late log phase culture (typically $\approx 4.0 \times 10^9$ cells total) were pelleted by centrifugation ($1500 \times g$, 10 minutes, 4°C), and then washed and treated as above to produce the P1 hydrogenosomal pellet. This was then resuspended to a final volume of 6ml in SHDI with 10% (v/v) iodixanol (Optiprep, Axis-Shield see p235). The suspension was then loaded on top of 10ml of SHDI/iodixanol gradient (20-40% Iodixanol) and overlain with 500 μl of SHDI. This suspension-gradient stack was then subjected to ultracentrifugation ($70,000 \times g$, 2 hours, 4°C) to density separate the organelles. After centrifugation the dark band corresponding to isolated hydrogenosomes (see Figure 4.3.1 p95) was extracted using a syringe and diluted tenfold in SH plus protease inhibitors (25 $\mu\text{g/ml}$ TLCK and 10 $\mu\text{g/ml}$ leupeptin). The extracted organelles were then pelleted by centrifugation ($8000 \times g$, 10 minutes, 4°C). The quantity of material was assayed by examining the absorbance at 280nm of resuspended organelles were heat treated with 0.6% SDS at 90°C for 5 minutes against a similarly treated BSA standard. Organelles were resuspended in SH plus protease inhibitors and glycerol (10%v/v, 25 $\mu\text{g/ml}$ TLCK and 10 $\mu\text{g/ml}$ leupeptin) before freezing at -80°C .

2.9 Hydrogenosomal manipulation techniques

2.9.1 Solubilisation of hydrogenosomal membranes

Thawed hydrogenosomes were collected by centrifugation (8000 x g, 5 minutes, 4°C), and resuspended in a solubilisation buffer (10% glycerol (v/v), 50mM HEPES, 5mM MgCl₂, 5mM DTT, 5mM EDTA, 1% protease inhibitor cocktail V (Merck) pH 7.4) with concentrations of NaCl between 250mM and 500mM, and detergents typically between 0.25% and 1%, nominal concentrations of detergents are shown below. Resuspended organelles were then rotated at 4°C for 90 minutes, before the lysate was separated from remaining organellar debris by centrifugation at 16,000 x g for 15 minutes.

| Detergent | CMC | | Concentration (%w/v) | Concentration as xCMC (xCMC) |
|--------------|------|--------|-------------------------|---------------------------------|
| | mM | (%w/v) | | |
| DDM* | 0.16 | 0.082 | 1.00 | 12.2 |
| | | | 0.50 | 6.1 |
| | | | 0.25 | 3.0 |
| Digitonin | 0.5 | 0.061 | 1.00 | 16.4 |
| | | | 0.50 | 8.2 |
| | | | 0.25 | 4.1 |
| SDS** | 8 | 0.23 | 1.00 | 4.3 |
| | | | 0.50 | 2.2 |
| | | | 0.25 | 1.1 |
| Triton X-100 | 0.24 | 0.015 | 1.00 | 66.7 |
| | | | 0.50 | 33.3 |
| | | | 0.25 | 16.7 |

*DDM: n-dodecyl β-D-maltoside, **SDS: Sodium dodecyl sulphate,

Table 2.4 Properties of the detergents used in this study

This study employs a number of different detergents to investigate *T. vaginalis* membrane proteins, the most commonly employed detergents are illustrated in the table above at their nominally used concentrations used within this study, these data are presented alongside the essential characteristics of these detergents, showing their critical micelle concentration (CMC) and the relation to the concentrations used practically.

2.9.2 Preparation of purified hydrogenosomal membranes

1mg of purified hydrogenosomes were pelleted by centrifugation (8,000 x g, 10 minutes, at 4°C) and were resuspended in 100 µl of 0.1M NaCO₃ with vigorous pipetting before being brought to a final volume of 1ml. The suspension was then rotated for 1 hour at 4°C to remove electrostatically interacting peripheral proteins before membraneous material was pelleted by ultracentrifugation at 140,000 x g for 30 minutes. Pelleted material was then gently washed to remove final traces of soluble material before the membrane pellet was resuspended in protein sample buffer. This method is also illustrated on p184.

2.9.3 Crosslinking of hydrogenosomal proteins

To covalently stabilise adjacent proteins chemical crosslinking was employed, in this investigation two DMSO soluble crosslinkers are used, DSP (Dithiobis[succinimidyl propionate], Lomant's reagent) and MBS (3-Maleimidobenzoyl-N-hydroxysuccinimide ester) which provide different crosslinking spans. (DSP: 12Å, MBS: 9.9Å). Hydrogenosomal membranes, or whole hydrogenosomes were resuspended in 200 µl/mg SH, crosslinking agents; DSP or MBS, dissolved in DMSO (to 20mM) were added to the suspension to a final crosslinker concentration of 0.1-1mM, and incubated on ice for 1hr. After the incubation excess crosslinker was quenched by the addition of 1M Tris, pH7.5, with continued incubation for a further 15 minutes on ice.

2.10 Protein import assays

Frozen or fresh organelles were resuspended in SH buffer to a concentration of 10mg/ml. Frozen cytosol was quickly thawed and for ATP-free conditions incubated with 1U/100µl of apyrase at 37°C. A minimal import assay consisted of 50 µl of either an ATP generating buffer: IB6 (1% w/v BSA, 275mM sucrose, 20mM MgCl₂, 100mM KCl, 11mM HEPES, 0.4mM sodium-succinate, 20mM sodium-pyruvate) or a neutral buffer: IB7 (1% w/v BSA, 275mM sucrose, 20mM MgCl₂, 100mM KCl), 20 µl of hydrogenosomal suspension, 20 µl of crude cytosol, and either 5 µl of SH buffer, or 2 µl of 0.1M ATP plus 3 µl of H₂O. In reactions supplemented with ATP, Na-ATP (Sigma) was resuspended in 10mM Tris (pH 7.5) at a concentration of 0.1M. reactions mixtures were then briefly mixed before incubation at 37°C for five minutes before the addition of 50ng of preprotein dissolved in 8M Urea followed by further mixing. Incubation was

then allowed to proceed for a further 30 minutes, before import reactions were halted by the addition of 1ml of SH buffer, and pelleting of the organelles by centrifugation (8,000 x *g* for 5 minutes at 4°C). Organelles were resuspended in 200 µl of SH buffer, with trypsin (50 µg), or trypsin and Triton X-100 (0.1% v/v) supplementation to assess whether the precursor protein was protected from protease digestion by virtue of association with organelles or import into organelles. Trypsin containing reactions were halted by the addition of 2 µl of 20mg/ml Soy Bean Trypsin Inhibitor. Assays were then collected by centrifugation at 8,000 x *g* for 5 minutes at 4°C, with the organellar pellet being resuspended in 200 µl SH, and the lysate being retained in experiments containing Triton. The protein content of the assays was then precipitated by TCA as previously described.

2.11 Immunological techniques

2.11.1 Antibodies

Antibodies used in this project are shown in the table below;

| Antibody | Use | Animal | Concentration |
|--|--|--------|----------------------|
| Primary antibodies | | | |
| Monoclonal Anti-HA, Clone HA-7 (Sigma) | Confocal microscopy, western blotting, Co-IP | Mouse | 1:2,000 to 1:20,000 |
| Anti-BIP (Novagen) | Confocal microscopy | Rabbit | 1:5,000 |
| Anti-Hmp43 (<i>T. vaginalis</i> protein), (Thermosystems) | Western blotting | Rabbit | 1:250 -1:500 |
| Secondary antibodies | | | |
| Polyclonal swine anti rabbit (Novagen) | Western blotting | Swine | 1:10,000 |
| Goat anti mouse (Sigma) | Western blotting | Goat | 1:10,000 to 1:25,000 |
| Conjugates | | | |

| | | | |
|----------------------------------|-------------------------|-------|---------|
| Anti-His, HRP conjugate (Qiagen) | Western blotting, Co-IP | Mouse | 1:5,000 |
|----------------------------------|-------------------------|-------|---------|

Table 2.5 Antibodies used within this study

Antibodies were used for the various aspects of this study, but most importantly for detection in both microscopy and biochemical assays, and for protein purification and enrichment. The table above illustrates the antibodies used in this study, showing their epitopes, sources and suppliers. Nominal concentrations are show indicated, Co-immunoprecipitation typically used large quantities of antibody, and the concentrations employed are mentioned in Methods 2.11.3.

2.11.2 Immunological detection of proteins transferred to PVDF membranes

PVDF membranes were wetted in methanol before being saturated in (Tris buffered saline) TBS +0.1%(v/v) Tween 20 (TBST). Wetted membranes were then blocked in TBST +5%(w/v) milk powder for hour at room temperature with rocking, the membranes were then washed 5 times with TBST for 5 minutes at room temperature with rocking. Incubation with primary antibody was performed at an appropriate dilution (Table 2.5) for 1 hr at room temperature with rocking. The membrane was subsequently washed with TBST as before and the secondary antibody diluted to 1:20,000 and allowed to incubate for 1hr at room temperature with rocking. Before ECL the membrane was washed with TBST as previously, before premixed EZ-ECL chemi-luminescent substrate was added to the membrane. Imaging of PVDF membranes was obtained using a Fujifilm LAS-100 and processed using the Fujifilm LAS-1000 software.

2.11.3 Co-immuno precipitation of HA-tagged proteins

Detergent solubilised hydrogenosomes were incubated overnight at 4°C with 2 µl of monoclonal anti-HA antibody (Clone HA-7, Sigma) with rotation. Immuno detected protein was recovered by incubation with 20 µl Protein A sepharose beads, which were resuspended in 1% (v/v) Triton X-100, 0.5% (w/v) BSA, 0.1M Tris, 0.3M NaCl, 1mM β-mercaptoethanol, 1%(v/v) Protease Inhibitor cocktail V (Merck), prior to incubation. Beads were allowed to incubate with the cleared lysate for 1 hour at room temperature or overnight at 4°C with rotation before being collected by centrifugation (3,000 x g, 3 minutes at 4°C). To remove non specific protein binding the beads were washed 4 times with 1ml of 0.1% (v/v) Triton X-100, 0.1M Tris, 0.3M NaCl, 1%(v/v) Protease Inhibitor cocktail V followed by centrifugation (3,000 x g, 3 minutes at 4°C) to recollect the beads. Bound proteins were eluted by the addition of 2x bead volume of 0.1M glycine at pH 2.8 with gentle mixing for 5 minutes, the eluate was extracted by pelleting the beads by centrifugation (3,000 x g, 3 minutes at 4°C).

2.12 Confocal Microscopy

2.12.1 Confocal system

Slides were analysed on a Zeiss LSM 710 confocal laser scanning microscope, using a 63x oil-immersion objective.

2.12.2 Preparation of *T. vaginalis* cells

Late log stage cells from transformant cultures were separated from media by centrifugation at 1500 x g for 10 minutes at 4°C. Pelleted cells were resuspended to a density of 1×10^7 /ml in phosphate buffered saline (PBS) and transferred to silane covered microscope slides (Sigma) and left to adhere for 30 minutes at room temperature. Non-adhered cells were washed from the slide with PBS and remaining cells fixed and permeabilised with 4% w/v paraformaldehyde, 0.1% v/v Triton X-100 for 20 minutes at room temperature.

2.12.3 Immunodetection of HA-tagged proteins

Permeabilised cells were washed with PBS before being blocked in supplemented PBS (+0.25% (w/v) BSA +0.25% (w/v) fish scale gelatin- PBSBG). The cells were then incubated in primary anti-HA antibody at a concentration of 1:2500 in PBSBG for one hour at room temperature. The slides were then washed with PBS twice before incubation of the secondary Alexafluor 488 coupled secondary in PBSBG diluted to 1:1000 for one hour.

2.12.4 Nuclear staining with propidium iodide

To stain the nuclei, slides were then washed twice with PBS and then incubated with PBS +RNAase (100 μ g/ml) at 37° for 20 minutes (to degrade any cellular RNA), washed with PBS before being incubated with PBS + 3.3 μ g/ml Propidium Iodide (PI) for 5 minutes, excess PI was removed with PBS washes before slides were mounted in 50% glycerol(v/v)/PBS.

2.12.5 Tracking agents

Golgi staining of fixed cells was achieved by incubation of live cells with NBD C6-ceramide (Novagen) in PBS at a concentration 1mM for a period of 30 minutes, cells were then washed twice with PBS before further labeling.

LysoTracker (DND 99, Novagen) was entrained into live cells suspended in PBS at a concentration of 50nM for a period of 30 minutes. Cells were then washed twice with PBS, before application to the slide.

These markers rely on the deposition of fat soluble markers within the membrane bilayer, during investigation Triton X-100/paraformaldehyde was found to adversely effect signal, and so slides with these trackers were simply fixed in 4% w/v paraformaldehyde.

2.12.6 Imaging software

Carl Zeiss Zen 2009 Light Edition (ZEN LE)

(<http://www.zeiss.com/>)

This was the principal software used in conjunction with the data obtained from the Zeiss LSM 710 microscope. Most of the high resolution confocal microscopy images seen in this work have been processed in this software. This software was also used for voxel based 3D reconstruction of images.

Carl Zeiss LSM Image Browser software

(<http://www.zeiss.com/>)

This software was used to examine Zeiss LSM images, and to derive measurements of features seen within the Carl Zeiss Zen 2009 Light Edition software.

ImageJ

(Abramoff & Magelhaes 2004) (<http://rsb.info.nih.gov/ij/index.html>)

Images on the Leica confocal microscope were analyzed using the ImageJ program, in particular for the derivation of the scale and resolution criteria for images off of this microscope.

bioView3D

(Kvilekval et al. 2009) (<http://www.bioimage.ucsb.edu/downloads/BioView3D>)

Advanced threshold analysis of Z-stack data was performed in bioView 3D, which proved to be an excellent tool for examining the 3D structure of Z-stack images.

2.13 Bioinformatic analysis tools

Further details on the implementation of bioinformatics facilities will be described in the relevant results section, but in brief:

BLAST

(<http://blast.ncbi.nlm.nih.gov/Blast.cgi>)

The BLAST algorithm was the principle tool used to detect sequence homologues in other species, it has also been used to align selected protein sequences. The modification to its parameters allow for specialized searches, for example to screen particular clades for sequence homologues.

Pfam

(Bateman et al. 2002) (<http://pfam.sanger.ac.uk/>)

The Pfam repository contains information related to hundreds of protein families, and information on protein domains. It has been used in this project to identify protein domains from query sequences

PHYLP

(Retief 2000) (<http://evolution.genetics.washington.edu/phylip.html>)

The Phylip tool suite was used in conjunction with other tools to do specific searches, alignments, or consensus tree generation. Whilst other programs incorporate features of this toolset, the original programs are contained within this package

BioEdit

(Hall 1999) (<http://www.mbio.ncsu.edu/bioedit/page2.html>)

The BioEdit was used to collate information from other searches, and to visually inspect and edit alignments. This program also include phylogenetic and other analysis tools, making it an excellent general workspace

CDD

(Marchler-Bauer et al. 2011)

Domain architectures returned in BLAST were examined in CDD to briefly examine whether query and result protein sequences share similar domain organisation as well as homology.

2.13.1 Phylogenetic analysis

CLUSTALX

(Jeanmougin et al. 1998) (<ftp://ftp.ebi.ac.uk/pub/software/clustalw2>)

The Clustal X program was used to derive alignments, and to produce phylogenetic trees

Dendroscope

(Huson et al. 2007) (<http://ab.inf.uni-tuebingen.de/software/dendroscope/>)

Dendroscope was used to image, and manipulate phylogenetic trees, and allows for user edition of the style and format.

2.1.1 Membrane topology

TopPred 2

(Claros & von Heijne 1994) (<http://www.sbc.su.se/~erikw/toppred2/>)

This program was used to predict transmembrane helices in some predicted *T. vaginalis* translocase homologues, prediction is made by window searching the primary sequence for regions of hydrophobicity.

Pred-TMBB

(Bagos et al. 2004) (<http://biophysics.biol.uoa.gr/PRED-TMBB/>)

The Pred-TMBB program was used to examine the primary sequence of β -barrel candidates, this tool uses a HMM algorithm to identify β -strand structure. This program was also used to score the probability of the proteins, as well as to determine strand count within the β -barrel.

3 Approach to practical work

To fulfill the objective set in the investigation- to characterize the *T. vaginalis* hydrogenosomal preprotein import system, a variety of techniques were employed to probe different parts of the translocation machinery. These approaches are first introduced here to show how they integrate in a workflow designed to meet the aims set in the introduction.

Principally the experimental work involves two different practical routes. One route attempted to find the translocon directly. This “top down” approach attempted to reconstruct a model system *in vitro* in combination with a molecular probe, a hydrogenosomal preprotein, to interact with the translocation machinery.

This approach has some significant advantages, as it requires no foreknowledge of the translocase proteins and has been used successfully in mitochondrial research (Glick & Schatz 1991; Cuezva et al. 1993; Schwarz & Neupert 1994), this method has also been redeveloped for the hydrogenosome (Bradley 1997) and was used to determine the essential characteristics of preprotein import. This study aimed to reproduce these previous results, and build upon them by characterizing components of the translocase system (Chapter 4).

This experimental method required the development of purification strategies to obtain isolated hydrogenosomes and cytosol from *T. vaginalis* C1 to create the *in vitro* protein import reactions. In addition a hydrogenosomal preprotein was synthesized and purified in an *E. coli* expression system. This exogenous protein had to be detectable with both immunological detection techniques using an in-frame poly-His tag, but also by radiological detection through the synthesis of ³⁵S labeled protein.

The preprotein import assay is described in Methods 2.10 (p75), and uses SDS PAGE to examine the sensitivity of the preprotein probe to trypsin. This assay was then developed to probe the aspects of the hydrogenosomal translocase system.

The second methodology used in the investigation involved a “bottom up” approach whereby candidate proteins were selected on the results of a bioinformatic analysis (Chapter 5), and were then investigated using a series of microscopy (Chapter 6) and molecular biology techniques (Chapter 7), using transformed *T. vaginalis* strains. Following localisation investigation into the complexes to which candidate proteins formed were investigated by a variety of techniques, ultimately developing methods to purify these for further analysis.

The latter approach has the advantage over the “top down” methodology as it allows the characterization of proteins and complexes which may not be associated with the initial preprotein binding, or for the processing of preproteins in distinct translocation pathways, for example insertion into the outer membrane. This approach also allows the investigation of transiently associating peripheral proteins, and to more broadly probe the scope of the general translocase system.

| “Top down” approach (Chapter 4) | “Bottom up” approach (Chapters 5, 6, 7) |
|---|---|
| Generation of plasmids encoding hydrogenosomal matrix pre-proteins | Generation of plasmids each encoding putative hydrogenosomal translocase proteins |
| Purification from an <i>E. coli</i> expression system of recombinant pre-protein | Transformation of <i>T. vaginalis</i> C1 to generate strains with endogenously expressed candidate translocase proteins |
| Purified hydrogenosomes from the parent strain <i>T. vaginalis</i> C1 are collected | Hydrogenosomes from each transformant strain are purified and analyzed |
| Import assay uses purified hydrogenosomes and cytosol from <i>T. vaginalis</i> C1 | Transformant cells can be fixed for microscopy to determine the localisation of candidate proteins |

Both approaches then use immunoprecipitation and electrophoretic techniques to examine protein complexes.

Figure 3.1 Practical aspects of the two experimental approaches

The essential differences and similarities of the two different methodologies are shown in the table above. Both approaches require plasmids expressing hydrogenosomal proteins, the top down methodology exogenously expresses these in *E. coli* whilst the bottom up approach creates recombinant strains of *T. vaginalis*. The top down approach uses pre-protein import assays to probe the nature of the translocon, where as the bottom up approach creates the opportunities to investigate the hydrogenosome directly. Both approaches are subsequently analysed by standard biochemical techniques.

The essential features of these two approaches are shown in Figure 3.2. Both approaches are complementary and explore the preprotein translocase system of the hydrogenosome. In combination these techniques widen the scope for discovery within this investigation, and encompass the greater part of the translocase system, these methods also allow corroboration of data by multiple approaches. In addition a multi-methodology approach is more flexible when elements of the workflow prove to be intractable, as alternate routes to investigation are available.

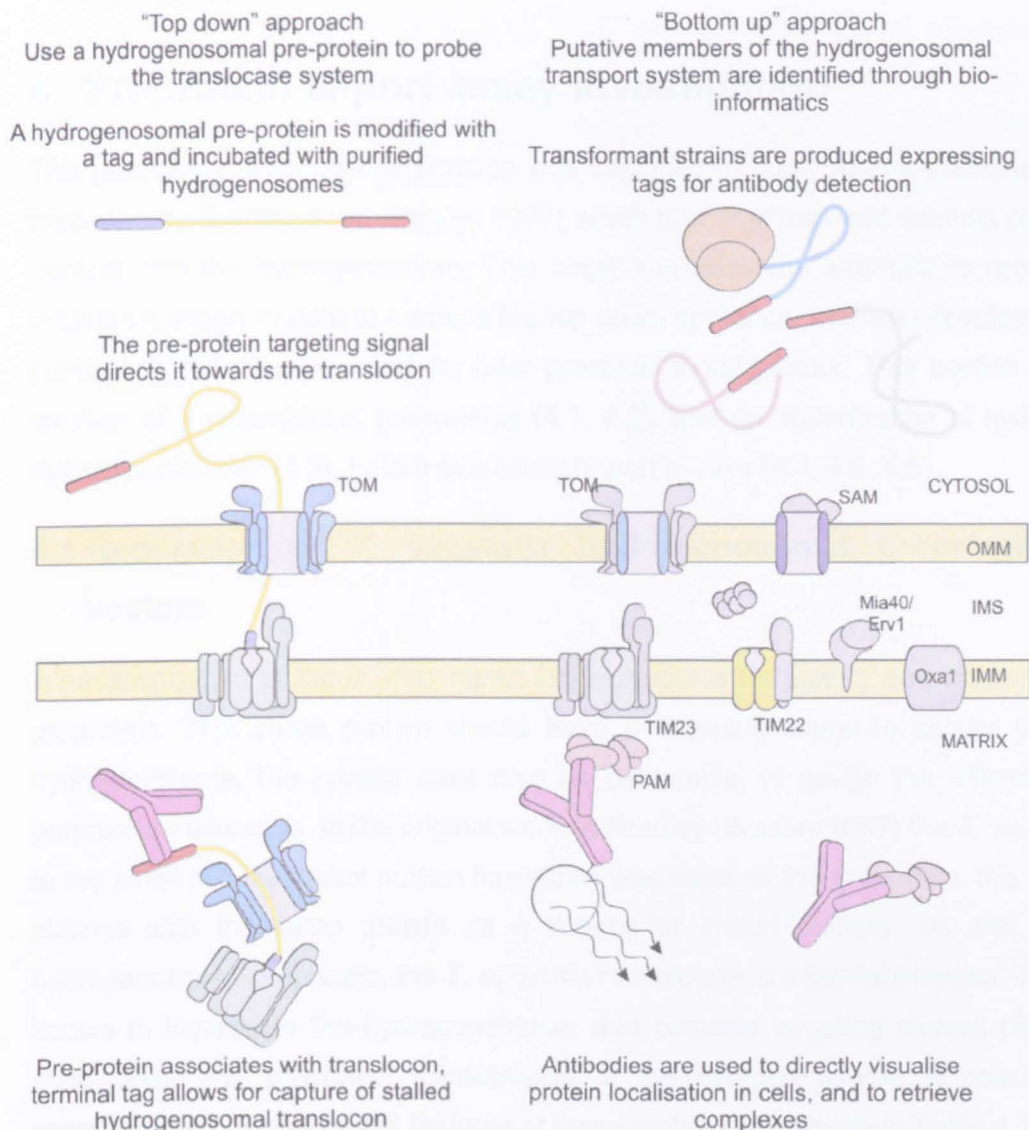


Figure 3.2 Workflow methodology for exploring the preprotein import system of *T. vaginalis*

Two separate approaches are taken in the practical investigation of the import machinery of the *T. vaginalis* hydrogenosome. The first thematic workflow takes a top down approach to examine the import system. Native hydrogenosomal proteins are exogenously expressed in an *E. coli* expression system with a recombinant C-terminal His-tag (red). Purified preprotein is then incubated with purified *T. vaginalis* hydrogenosomes, and conditions are used to generate stalled translocons containing the recombinant preprotein. Experimental approaches are then used to extract information about the translocon. The second approach uses a bottom up perspective whereby putative members of pre-protein translocase complexes are identified through bioinformatics, and transformants are generated with tagged hydrogenosomal proteins. Immunological methods are then used to probe the localisation of the recombinant protein and extract complexes for further analysis.

4 Preprotein import assay investigation

This part of the practical investigation was designed to build upon a preprotein import system produced by Bradley et al (Bradley 1997) which first characterized aspects of how preproteins interact with the hydrogenosome. This section outlines the attempts to recreate the *in vitro* preprotein import system to address the top down approach, and also develop hydrogenosomal purification techniques critical for later practical investigations. This section first address the creation of a recombinant preproteins (4.1, 4.2), and the optimization of hydrogenosome and cytosol purification (4.3), before examining import assays (4.4, 4.5, 4.6).

4.1 Generation of *T. vaginalis* hydrogenosomal preprotein expression vectors

A necessary part of the *in vitro* import assay model is the use of a hydrogenosomal precursor preprotein. This probe protein should have a targeting signal to deliver the protein to the hydrogenosome, the protein must also be detectable, to gauge the efficacy of binding and preprotein maturation. In the original work by Bradley (Bradley 1997) the *T. vaginalis* homologue to the small soluble matrix protein ferredoxin was used as the preprotein, this study recreates a plasmid with the same protein as a control for import efficacy, as well as an additional hydrogenosomal preprotein, the *T. vaginalis* homologue to adenylate kinase. These proteins are known to localise to the hydrogenosome, and possess targeting signals (Bradley 1997) and must utilise the preprotein translocases of the hydrogenosome to reach their functional compartments. The essential features of these proteins are shown in Table 4.1.

Sequences for the preproteins were identified in the genome sequenced *T. vaginalis* G3, (Carlton et al. 2007). These hydrogenosomal genes are intron-less, and are able to be produced directly from gDNA. Primers were designed based upon this genome data (Table 4.1) and DNA containing the complete preprotein sequences were produced by PCR from *T. vaginalis* G3 gDNA (Methods p66). The products of the PCR reactions can be seen in Figure 4.1.1 (adenylate kinase) and Figure 4.1.2 (ferredoxin). Whilst a single clear product can be seen in the gDNA PCR for ferredoxin, a doublet species can be seen for adenylate kinase. Only material from the full size fragment was used for ligation into the recipient vector, pET26b (Novagen).

Figure 4.1.1 and Figure 4.1.2 also show the results of miniprep digestion from kanamycin resistant transformants carrying insert ligated plasmid. The insert fragment is seen recovered (if weakly for ferredoxin). Sequencing PCR using BigDyes (Described in Methods2.6.5.1 p68) was used to validate insert fidelity and orientation. The constructs were maintained in an *E.coli* DH5 α strain, and expression in a BL21 strain transformed from purified DH5 α plasmid.

| Name | <i>T. vaginalis</i> protein | GI: Number | Length |
|----------------------------|--|------------|--------------|
| pTVAGAK | Adenylate kinase | GI:463922 | 687bp, 229aa |
| Parent vector | pET 26b | | |
| Forward primer (AKFOR) | 5'-GGCCTT[CATATG]CTCAGTACATTAGCTAAG-3' | | |
| Reverse primer (AKREV) | 5'-TTGGCC[GGATCC]CTGGAGGGCAGCATGAAGCTTC-3' | | |
| Restriction enzymes | NdeI, XhoI | | |
| Features | C-terminal His-tag, kanamycin resistance, IPTG induction | | |
| pTVAGFER | Ferredoxin | GI:162509 | 300bp, 100aa |
| Parent vector | pET 26b | | |
| Forward primer (FDXFOR) | 5'-GGCCTT[CATATG]CTCTCTCAAGTTTGCCGC-3' | | |
| Reverse primer (FDXREV) | 5'-TTGGCC[GGATCC]GAGCTCGAAAACAGCACCATCG-3' | | |
| Restriction enzymes | NdeI, XhoI | | |
| Features | C-terminal His-tag, kanamycin resistance, IPTG induction | | |

Table 4.1 Essential features of the pre-protein expression constructs used in this study

Two preprotein constructs were created for the import study, pTVAGAK expressing hydrogenosomal adenylate kinase, and pTVAGFER containing the *T. vaginalis* ferredoxin protein which has been previously shown to target to the hydrogenosome (Bradley 1997). The table above shows the essential features of the plasmids. The primers used to amplify the sequences from gDNA are shown, as are the added restriction sites (square brackets). Features of the plasmid include a C-terminal poly-His tag to purify recombinant protein, and induction by LacI with lactose analogs (IPTG). Selection for transformed cells was by kanamycin resistance.

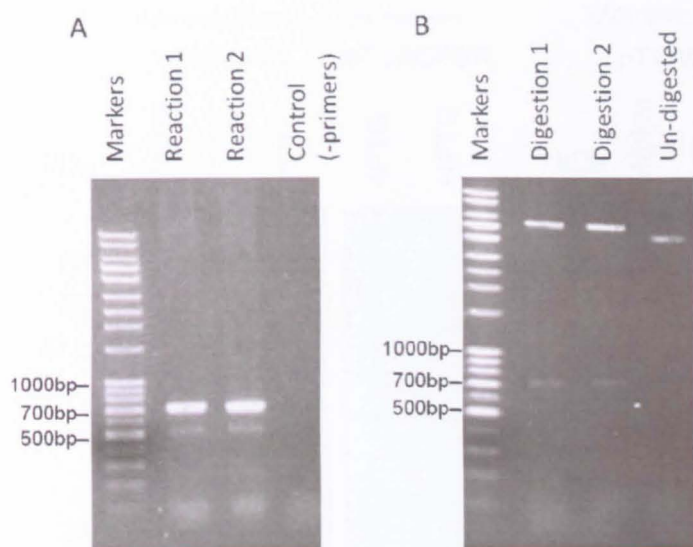


Figure 4.1.1 Preparation of the pTVAGAK plasmid

T. vaginalis G3 gDNA was used as a template for a PCR reaction to extract a fragment encompassing the intron-less *T. vaginalis* adenylate kinase gene, the results of this reaction are shown in A. A bright species at the expected size (687 bp) is seen corresponding to the coding sequence for the protein precursor. In addition a fainter band is seen below. The larger fragment was excised and ligated into a doubly restricted pET26b to create pTVAGAK. Ligated plasmid was used to transform *E. coli* DH5 α cells, and mini-prepped plasmid was digested with the same enzymes used to create the ligation sites (panel B), where the original fragment is recovered.

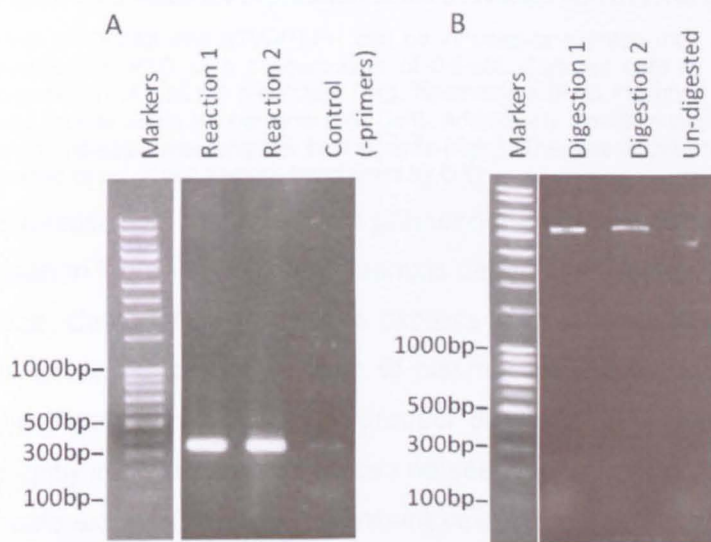


Figure 4.1.2 Preparation of the pTVAGFER plasmid

Preparation of the pTVAGFER is shown in this figure, similar to Figure 4.1.1, the results of a gDNA PCR on *T. vaginalis* G3 gDNA is shown in A, only a single fragment for ferredoxin was produced and was of the expected size (300 bp). This fragment was used to generate the pTVAGFER plasmid, whose digestion with NdeI and XhoI is shown in B, the ferredoxin band is seen faintly at 300bp.

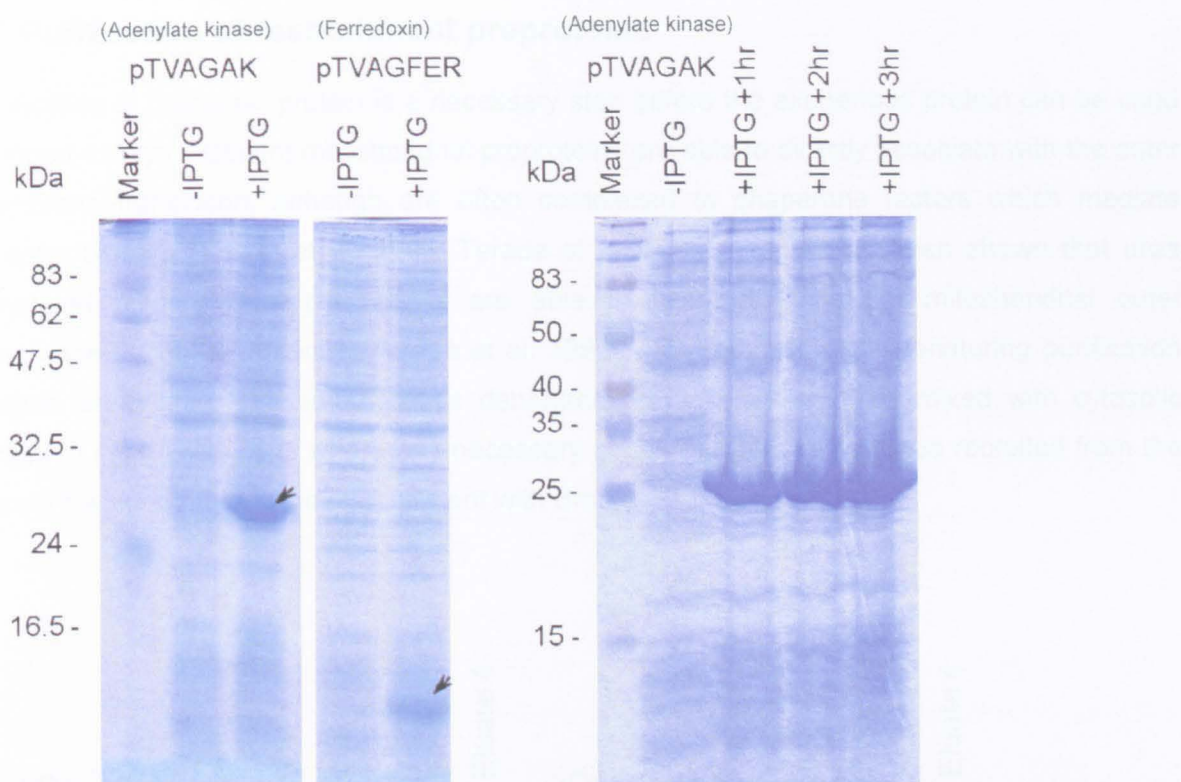


Figure 4.1.3 Inducible expression of the pTVAGAK and pTVAGFER expression constructs

Both pTVAGAK and pTVAGFER can be induced to express their inserted proteins through the LacI operator, by addition of IPTG at a concentration of 0.5mM. Cultured cells in the absence of IPTG do not show significant expression of plasmid preprotein (left), however on IPTG induction (for 1hr) sizable quantities of preprotein rapidly accumulate within the bacterial cells (left). Additionally induction yield appears to saturated within one hour, with only slight increase over lengthier time periods (right). The gels in this image were produced from whole lysed cells, and loaded to equal cell density determined by O.D.

Expression of the preprotein precursors was demonstrated in BL21 by induction with IPTG as seen in Figure 4.1.3. Both plasmids demonstrate the ability to express a protein of the expected size. Cells did not appear to express a sizable quantity of precursor in the absence of IPTG, however induced expression of plasmid protein was seen when IPTG was introduced, with a significant accumulation of product obtained after one hour. Lengthier expression times did slightly improve yield which can be seen in Figure 4.1.3 for pTVAGAK. From these results large scale expressions of transformant cells were performed to collect whole cell pellets for protein purification.

4.2 Purification of recombinant preproteins

Purification of precursor protein is a necessary step before the exogenous protein can be used in import assays. Nascent mitochondrial preproteins are able to directly associate with the outer membrane translocon, although are often complexed to chaperone factors which mediate receptor binding (Young et al. 2003; Terada et al. 1996). It has also been shown that urea denatured mitochondrial preproteins are able to associate with the mitochondrial outer membrane translocon directly (Terada et al. 1996). Considering this, a denaturing purification method using urea was employed as denatured preproteins would be mixed with cytosolic factors in import reactions where any necessary chaperone factors could be recruited from the cytosol fraction (this method is consistent with Bradley).

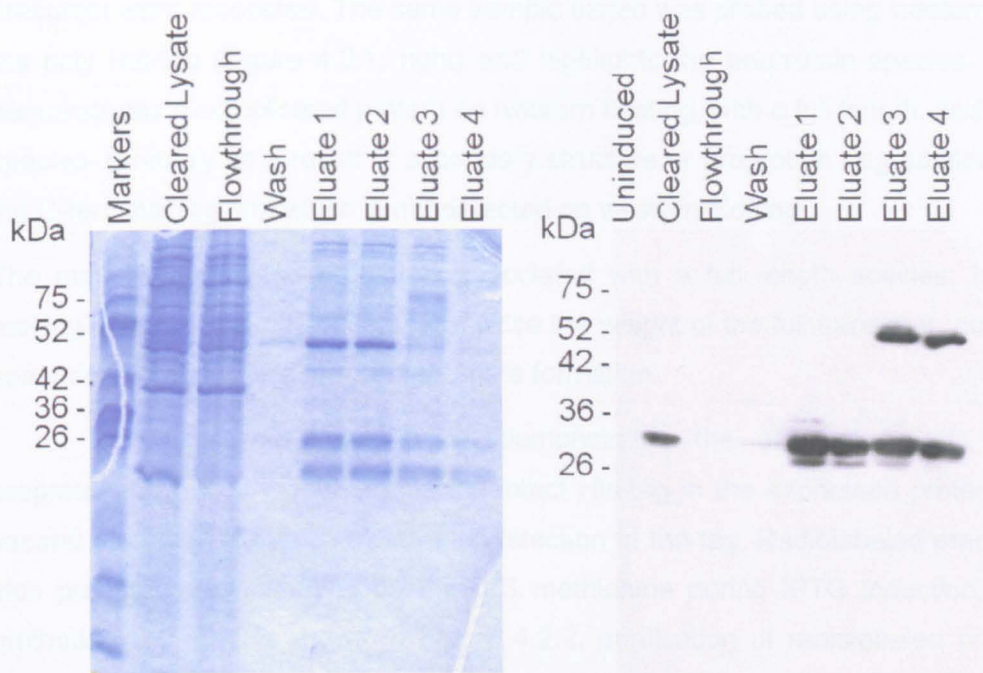


Figure 4.2.1 Purification of the adenylate kinase preprotein

Samples from a purification series from pTVAGAK are shown in the above figure, identical gels were run with samples of cleared lysate before and after nickel bead incubation (flowthrough), a sample of the wash eluate was run to ascertain precursor loss during bead washing, and a sample from a series of elutions with a low-pH, high-imidazole buffer. The left hand gel has been stained and indicates the presence of a full size precursor in the cleared lysate, as well as in the eluate samples. Western detection using the poly-His tag, highlights the same species at ~26kDa, in addition to a possible dimeric/aggregate form at ~52kDa.

Both constructs pTVAGAK and pTVAGFER express hydrogenosomal preproteins with a C-terminal poly His-tag (His₆), enabling purification by nickel bead purification. Transformant BL21 cells were induced with IPTG for 1 hr to accumulate expressed protein before being

pelleted and frozen. These pellets were then lysed in a high urea lysis buffer as described in Methods 2.7.8 (p72) this lysate was subsequently cleared by centrifugation.

Cleared lysate was incubated with nickel coated beads to bind His-tagged precursor preprotein, before being pelleted and washed, before preprotein elution using competitive displacement of the poly-His tag by imidazole under increasingly acidic conditions, where protonation of histidine residues acts to reduce competition. Figure 4.2.1 illustrates the purification of non-radiolabeled adenylate kinase using this technique.

Adenylate kinase was clearly expressed in the cleared lysate, and the lysate recovered after bead incubation (flow-through) is seen to be deficient in preprotein. Some small trace amount of precursor is lost during washing, however upon elution significant quantities of precursor were recovered. The same sample series was probed using western blotting to detect the poly His-tag (Figure 4.2.1, right) and highlights the preprotein species. Adenylate kinase demonstrates a complicated pattern on western blotting, with a full length, and apparently lighter species- probably as a result of secondary structure or preprotein degradation (possibly loss of the C-terminal region), which is not detected on western blotting.

The majority of purified protein is associated with a full length species. In addition a high molecular weight band is observed at twice the weight of the full transcript, potentially indicating some degree of stable dimer or aggregate formation.

The nickel bead purification demonstrates the ability to purify and concentrate preprotein, and also demonstrates the intact His-tag in the expressed protein, and provides a second means of detection by immunodetection of the tag. Radiolabeled precursor protein was also prepared by incubating cells in ^{35}S methionine during IPTG induction, described in the methods (p71) and is shown in Figure 4.2.2, purification of radiolabeled preprotein uses the same nickel-bead purification method as employed with the unlabelled protein. Nickel bead purification efficiently recovers protein from cell lysate, and depletes radiolabeled product from the flowthrough. Eluates exhibit a similar pattern to western-detected preprotein, but the lower molecular weight species are also detected, as detection does not require poly-His tag recognition for signal.

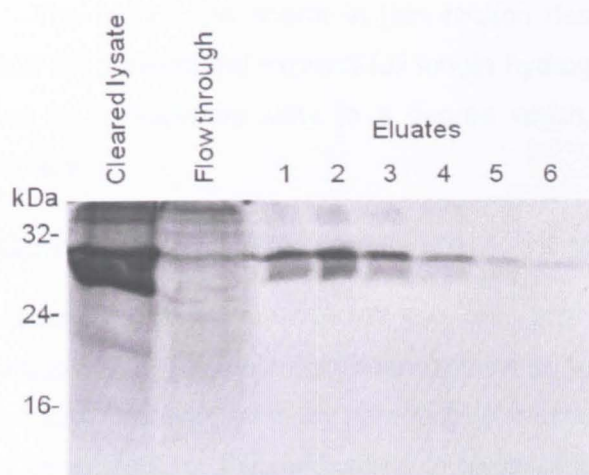


Figure 4.2.2 Purification of ^{35}S radiolabeled adenylate kinase from an *E. coli* expression system

The nickel bead purification strategy used for unlabeled His-tagged preprotein was also employed to purify radiolabeled protein from pelleted cells incubated with ^{35}S methionine. This figure shows radiolabeled adenylate kinase preprotein. Samples of flow through and cleared lysate are loaded (10 μl in 50 μl) alongside successive eluates of the bead bed (1 μl in 50 μl).

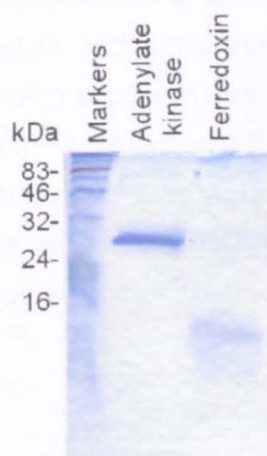


Figure 4.2.3 Preprotein purification using a centrifugal size concentrator

Protein contaminants remaining in His-tag purification eluates were specifically eliminated using a centrifugal size separation column. The purified preproteins adenylate kinase and ferredoxin are shown in the figure above. A centrifugal concentrator with a 25kDa cut-off was used to hold back higher molecular weight contaminants from the centrifugal eluate. Combined with the His-tag purification this strategy was capable of producing of urea suspended preproteins of high concentration and purity.

Some remaining protein species are however carried across from the purification, which are seen at high molecular weight in Figure 4.2.1, fortunately all the significant contaminants were above 25kDa, larger than either preprotein protein. To refine the initial purification, eluates were passed through a centrifugal size concentrator with a 25kDa cutoff which was used to elute the desired precursor whilst holding back the high molecular weight contaminants and aggregates (Figure 4.2.3).

The techniques shown in this section demonstrate that the plasmids pTVAGAK and pTVAGFER contain and express full length hydrogenosomal preproteins, and that they can be purified in a denatured state to a degree which can be used within an import assay type experiment.

4.3 Optimisation of hydrogenosome preparation

The requirements for reproducing the preprotein import system *in vitro* also relies on the preparation of *T. vaginalis* hydrogenosomes and cytosol. Some prior work has been done to refine these products, and an approach is mentioned by Bradley et al (Bradley 1997) which builds upon work for the purification of glycosomes in trypanosomes (Opperdoes et al. 1984) and previous organellar separation to localise hydrolase activity in *Trichomonas* (Lockwood et al. 1988).

Initial attempts to repeat these purification strategies could not consistently reproduce hydrogenosomes of adequate quality and yield. Two initial methods of lysis were tried, cell disruption and sonication. Both of these methods tended towards extensive disruption of organellar structures, and either reduced yield, or increased protein degradation in our attempts. A gentler approach was then optimised for the lysis of *T. vaginalis* using simple needle passage, this method had the advantage of being fast, quickly assessed for degree of disruption, and un-reliant on particular pieces of equipment for reproducible results. Optimisation of this lysis method was settled upon as the desired disruption method within this investigation.

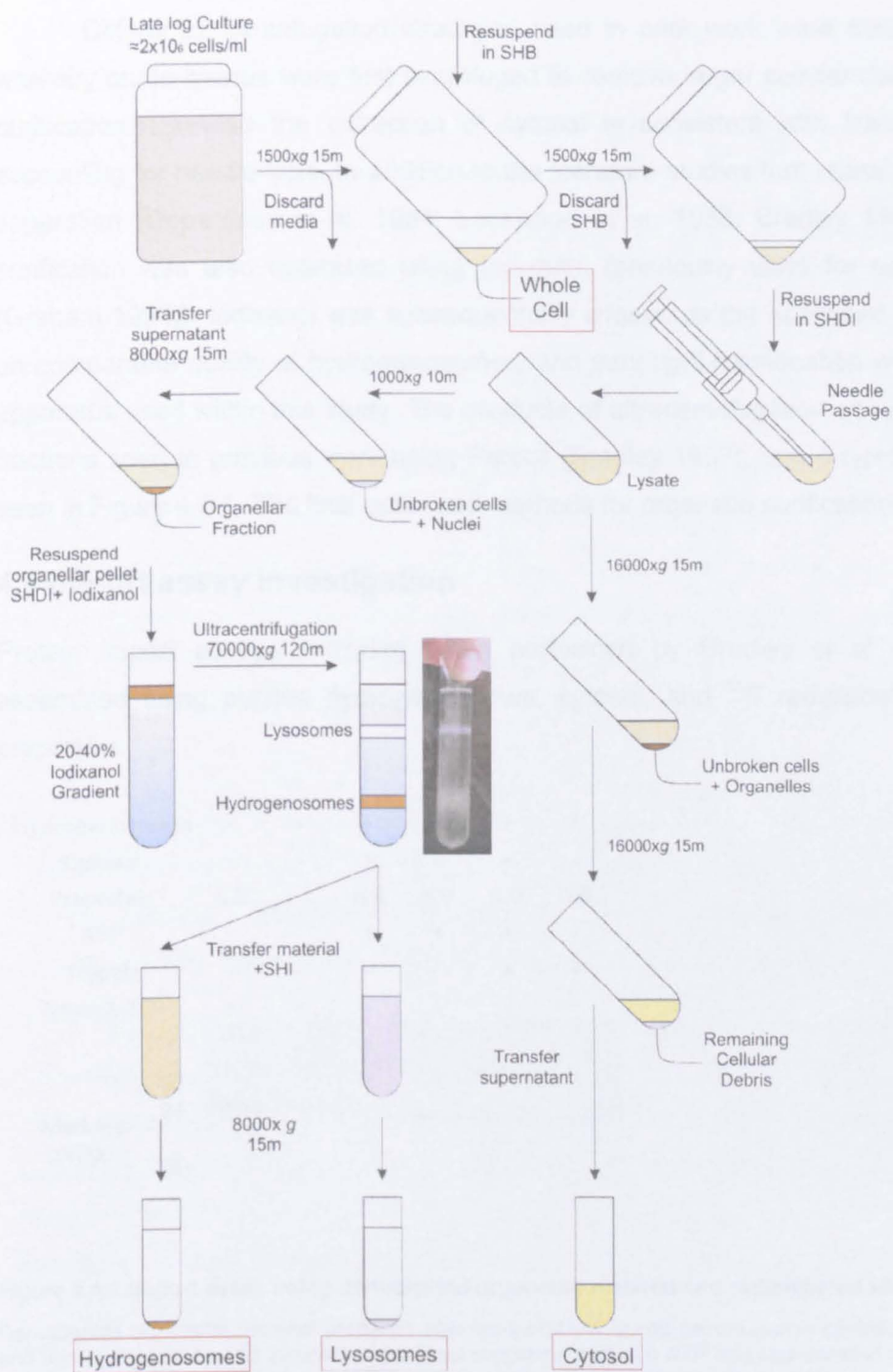


Figure 4.3.1 The experimental strategy for the derivation of the *T. vaginalis* cell fractions

This diagram illustrates the workflow for the derivation of cellular fractions from *T. vaginalis*. Dense cultures were pelleted and the cells washed (SHB, SHBI buffers), before the cells were lysed via needle passage. The crude lysate was subject to differential centrifugation to enrich an organellar fraction, which was then subject to density gradient ultracentrifugation to purify the organellar species. Whole cell and cytosol samples are also extracted from different points of the work flow. Full details of this process are given in methods p73

Differential centrifugation strategies used in prior work were maintained in this study, whereby crude lysates were first centrifuged to remove larger cellular debris before organellar purification. Likewise the extraction of cytosol is consistent with literature methods when accounting for needle lysis. In addition to the literature studies that utilised Percoll for organelle separation (Oppendoes et al. 1984; Lockwood et al, 1988; Bradley 1997), hydrogenosomal purification was also optimised using Iodixanol (previously used for organellar fractionation (Graham 1994)). Iodixanol was subsequently chosen as the optimised gradient agent based on comparable quality of hydrogenosomes, and very tight fractionation within the experimental apparatus used within this study. The products of ultracentrifugation compare favourably to the fractions seen in previous work using Percoll (Bradley 1997), and a typical purification can be seen in Figure 4.3.1. The final optimised methods for organelle purification can be seen on p72.

4.4 Import assay investigation

Protein import assays mirroring those performed by Bradley et al (Bradley 1997) were assembled using purified hydrogenosomes, cytosol, and ³⁵S radiolabelled adenylate kinase preprotein.

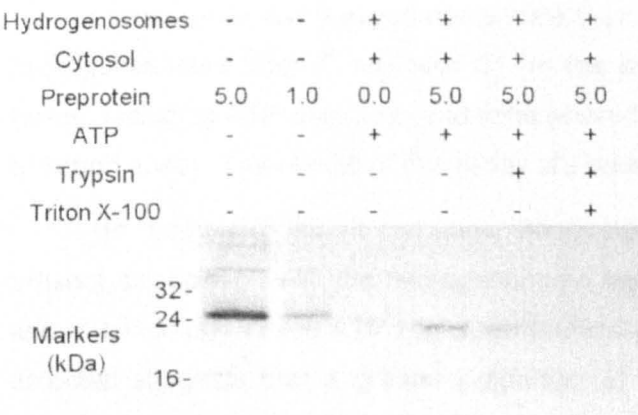


Figure 4.4.1 Import assay using denucleated organellar material and radiolabeled adenylate kinase

Denucleated organellar material prepared after lysis and low speed centrifugation containing a mixture of lysosomes and hydrogenosomes was incubated in cytosol supplemented with ATP and radiolabeled adenylate kinase precursor. After incubation organelles were pelleted by centrifugation and preprotein proteinase sensitivity was tested with trypsin. A negative control where Triton X-100 was used to disrupt the membranes is also shown, as well as a positive control with just radiolabeled precursor (50ng/μl) with the full dose used for the assay (5μl) and 20% of that dose (1μl)

Preprotein binding was initially trialled with denucleated organellar material, containing both hydrogenosomes and lysosomes with the radiolabeled precursor, the results of this assay are seen in Figure 4.4.1. The mixed organellar fraction was incubated with cytosol

supplemented with ATP and precursor before pelleting of the organellar fraction. Organelles were then resuspended in a sucrose/HEPES buffer, and subjected to trypsin digestion, a negative control was created where by organelles were solubilised using Triton X-100 to expose any internalised preprotein to the protease.

The results of this initial trial were encouraging and demonstrated there was some preprotein retention by the organelles after pelleting from suspension. Upon trypsin treatment some preprotein was lost, indicating that some of the organelle associating preprotein was exposed to trypsin, and with membrane disruption by Triton X-100, no signal was seen, indicating that all the preprotein had been exposed and digested. These results are consistent with a scenario where suspended preprotein in the cytosol is bound by the hydrogenosome, and that some of the preprotein was internalised or otherwise shielded from proteolysis, however on membrane disruption, detergent acted to expose any internalised or shielded protein to proteolytic attack. There is also some indication of a possible mass reduction with the precursor protein, however no doublet signal is discerned, and migration difference is slight. These data suggest that the cell lysis method is capable of generating organelles with the potential for import.

These preliminary experiments were then expanded and repeated with Iodixanol purified hydrogenosomes from *T. vaginalis* C1. In the expanded assay different ATP conditions were tested, including ATP depleted conditions where the cytosol had been treated with apyrase prior to import assay. The results of this assay are seen in Figure 4.4.2.

In “complete” import reactions; where hydrogenosomes, cytosol and preprotein are all present, association with the hydrogenosome happens in both ATP positive conditions (6, 8, 9) and in conditions where ATP has been artificially depleted using apyrase (10, 11). The signal detected suggests that a greater proportion of material is bound in ATP positive conditions (compare lanes 6 and 10), though quantification would be needed to support this claim. This phenomena would be consistent with published data (Bradley 1997).

On treatment with trypsin signal is reduced in both ATP conditions (lanes 8, 11). In ATP depleted reactions the signal is effectively eliminated in both Triton treated and untreated (lanes 11,12), which would suggest that the associated precursor was vulnerable to proteolytic attack without membrane disruption, consistent with the preprotein bound on the organelle periphery as opposed to being internalised. The requirement for ATP for preprotein protection is again consistent with literature.

The precursor protein is also seen to be retained even when it is incubated in the absence of hydrogenosomes or cytosol. For the protein to carry to detection in the "no hydrogenosome" reaction would suggest that the protein is binding to organelle fragments or factors that do pellet, signal however is very faint. Without cytosol, association either proceeds in the absence of cytosolic factors or that some factors survive hydrogenosomal purification, these results would reflect urea denatured precursor binding as seen for mitochondria (Terada et al. 1996). The results of these two series is unclear, however preprotein binding to isolated hydrogenosomes has already been published and suggest some ability for the hydrogenosome to bind some preproteins even in the absence of cytosolic cofactors (Bradley 1997).

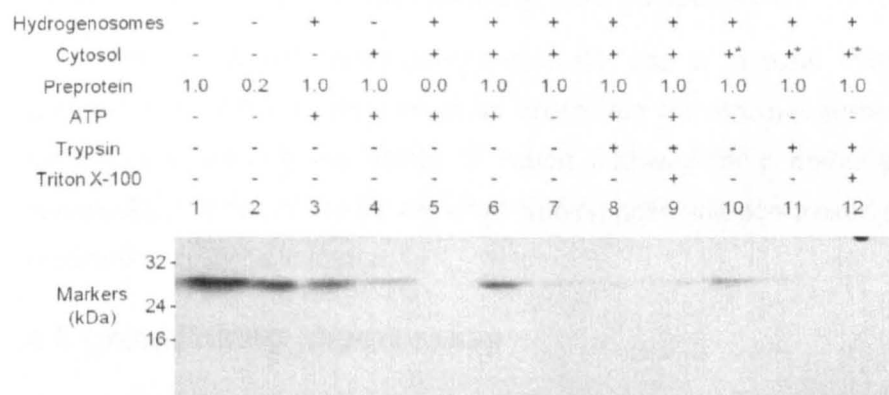


Figure 4.4.2 Preprotein association with purified hydrogenosomes in varying ATP conditions

A series of import assay reaction conditions were run with Iodixanol purified hydrogenosomes. Reactions were assembled with or without hydrogenosomes, cytosol (apyrase treated cytosol is indicated with *), preprotein (^{35}S radiolabeled adenylate kinase (50ng/ μl), loading is shown in μl) and ATP. Subsequent to incubation some reactions were treated with trypsin with or without Triton X-100.

Some protein binding is observed in reactions with untreated cytosol, but without supplemental ATP. This would suggest that preprotein association with the hydrogenosome does not necessarily require supplemental ATP. However the signal intensity is significantly weaker than the minus cytosol control. These data may suggest that supplemental ATP is required for best association when cytosolic factors are present, perhaps in a similar fashion to the ATP dependence required to transfer mitochondrial preproteins from Hsp90/70 onto mitochondrial receptors. (Young et al. 2003)

Signal is greatly diminished but not completely eliminated with Triton X-100 treatment, perhaps indicating solubilisation was incomplete. Masses in the complete import assay reactions appear marginally reduced however no doublet signal is observed and if there is a

genuine mass shift it is very hard to discern, and may not reflect a maturation of preprotein by signal sequence cleavage.

These experiments using the radiolabelled adenylate kinase precursor indicate that full import involving maturation of the precursor protein is not being achieved. However association with the hydrogenosome is consistently demonstrated as well as limited proteolytic protection in the presence of ATP, these results affirm what has been observed previously (Bradley 1997). Whilst we have been unable to reproduce the results of Bradley et al in that a reduced molecular weight species is not obtained, a partly functioning system whereby proteolytically protected preprotein has been generated in an ATP dependent fashion does demonstrate that even if full import is not being achieved then the precursor protein is still interacting with the hydrogenosome, and that this interaction is ATP dependent.

The “top-down” methodology can still use a specific hydrogenosomal interaction to uncover part of the hydrogenosomal preprotein translocase system. Whilst the scope of this technique to explore the whole of import pathway for a matrix preprotein is dependent on successful preprotein import, the initial binding sites and complexes can be still probed even with incomplete preprotein import.

4.5 Crosslinking import assay

The assays have shown that the precursor is stably bound to the hydrogenosome for the period of time that the organelles are exposed to the proteolytic agent, and that the stability of these complexes are on the order of tens of minutes at 4°C, however because the preprotein appears to only be part translocated the stability of the preprotein complex is probably inadequate to resolve any translocons by native PAGE type techniques.

To investigate the association with the hydrogenosome it was necessary to artificially stabilise these transient associations, and in this investigation chemical crosslinking was used to preserve the interaction partners. This approach allowed for the use of denaturing electrophoretic methods to probe complexes, as the crosslinking agent covalently links proteins.

| | | | | | | |
|----------------|-----|-----|-----|-----|-----|-----|
| Hydrogenosomes | - | + | - | + | + | + |
| Cytosol | - | - | + | + | + | + |
| Preprotein | 0.2 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 |
| ATP | - | + | + | + | + | + |
| Trypsin | - | - | - | - | + | + |
| Triton X-100 | - | - | - | - | - | + |
| | 1 | 2 | 3 | 4 | 5 | 6 |

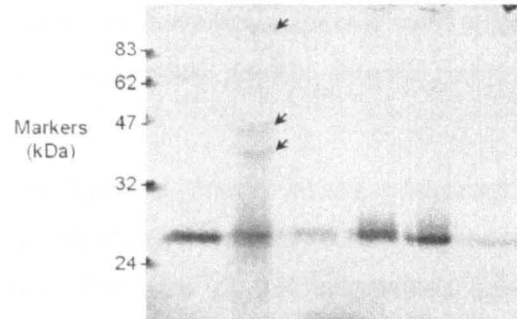


Figure 4.5.1 Preprotein chemically crosslinked during association with hydrogenosomes results in higher molecular weight species

An import assay was constructed to generate transiently associated adenylate kinase to its interacting factors. A representative 20% of total precursor is loaded on the left as a positive control, as well as control reactions missing either cytosol or hydrogenosomes. Import reactions in ATP supplemented conditions were also run, and chemically crosslinked in 0.5mM DSP after incubation, these hydrogenosomes are then challenged with trypsin and trypsin/Triton conditions to determine proteolytic protection.

In Figure 4.5.1 the results of an import assay which had been coupled post incubation with a chemical crosslinking step (DSP at 0.5mM,). This step would covalently link the preprotein to proximal hydrogenosomal proteins. Import reactions were then incubated with trypsin and trypsin/Triton X-100 to observe the effects of crosslinking on proteolytic digestion.

The reactions without hydrogenosomes or cytosol retained some precursor protein, and upon crosslinking generated novel species on electrophoresis representing covalently crosslinked species. Whilst a faint band is observed at $\approx 50\text{kDa}$ for the no hydrogenosome reaction, this signal is also observed in the precursor only control, and previously in the purification of the recombinant protein, and reflects a tendency of the preprotein to dimerise or aggregate. The reaction which was conducted in the absence of cytosol generated additional species which are not observed in the cytosol only reactions. Two molecular weight bands reflecting $\approx 10\%$ of the unlinked adenylate kinase signal can be observed at $\approx 38\text{kDa}$ and $\approx 47\text{kDa}$, additionally some signal intensity is observed at very high molecular weight ($\approx 100\text{kDa}$).

The results of the complete crosslinking reaction (that is including cytosol and hydrogenosomes) do not generate patterns similar to either cytosol, or hydrogenosome only reactions, but do exhibit some faint signal intensity at high molecular weight ($\approx 100\text{kDa}$; indicated with an arrow in Figure 4.5.1).

These results also suggest that DSP crosslinking does seem to positively affect protein stability in the presence of trypsin relative to previous assays (comparing lanes 4 and 5 in figure 4.5.1, as opposed to lanes 6 and 8 in Figure 4.4.2). However treatment does not protect the associating preprotein when the hydrogenosomes are treated with Triton. These data are consistent with a scenario whereby the preprotein was dynamically associating with a surface translocon, this dynamic movement between exposed and shielded surfaces is halted by covalent crosslinking and prevents shielded, possibly internalized preprotein from exposing itself to proteolytic cleavage.

The results of the crosslinked import assay encouragingly suggest that discrete complexes are formed, and some of the lower molecular weight complexes are resolved in the hydrogenosome only reaction. The size of the crosslinked species could arise from the formation of a bipartite complex containing adenylate kinase preprotein, and a 10-20kDa hydrogenosomal protein, this may represent an Preprotein-Receptor complex, as two preprotein receptors in the mitochondrial TOM complex are of a similar size (Tom20, 22). Other complexes might also exist, the composition of these are theorized in the discussion.

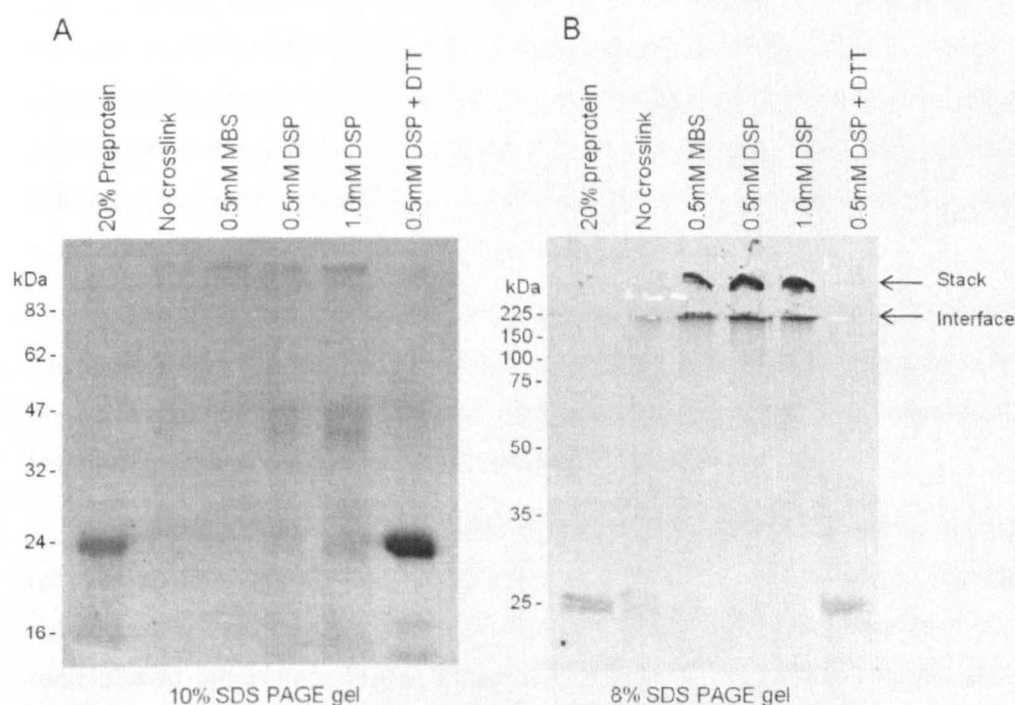


Figure 4.5.2 Optimization of chemical crosslinking for import assay reactions

Parallel import reactions with cytosol, and hydrogenosomes were incubated in supplemented ATP conditions in the presence of radiolabeled adenylate kinase preprotein. After initial incubation organelles were pelleted, and crosslinked with the crosslinking agents DSP and MBS. To explore different sized complexes, samples were run on 10% (A) and 8% gels (B).

To further investigate these results a complete reactions containing cytosol, hydrogenosomes, and radiolabeled protein in ATP supplemented conditions was tested in parallel under a range of crosslinking conditions with two crosslinking agents, the DTT reversible DSP and MBS. The results are shown in Figure 4.5.2.

The same doublet band pattern is seen repeated at $\approx 38\text{kDa}$ and $\approx 47\text{kDa}$ in the presence of crosslinking agents, but which is reversed on crosslinker cleavage with DTT, this is most clearly seen in the 10% SDS PAGE gel. Some additional high molecular signal is seen on the 10% gel, however resolution is not improved when duplicate samples are loaded onto an 8% gel, however the figure does clearly indicate that whilst crosslinking significantly increases the quantity of insoluble protein, seen trapped in the stack, some very high molecular complexes, or covalently linked aggregates are able to migrate into the resolving gel, however no structure to their distribution is observed.

4.6 Attempts to isolate preprotein translocon complexes

Whilst these electrophoresis methods had already begun to characterise complexes within in a mixed hydrogenosomal fraction, it was not possible to characterise these species beyond their masses. It was further investigated whether covalently crosslinked complexes containing preprotein could then be extracted from import reactions. This method would employ the same purification strategy already employed to purify the preprotein by utilising the C-terminal poly His-tag in denaturing conditions to chelate nickel. This technique would enable promising candidates to be identified by mass spectrometry.

The obstacles to this approach are that there are a small and finite number of sites on the hydrogenosome that the preprotein can bind to, and site saturation could be a problem with only a few nanograms of protein per mg of hydrogenosome. With effective crosslinking, only a few nanograms of protein would be available to purify.

Attempts to recover adenylate kinase from the standard import reactions were unable to recover a detectable yield. To improve the amount of precursor available reactions were subsequently scaled to 1mg (by protein content) of hydrogenosomes using 250ng of radiolabeled preprotein. Trace amount of monomeric adenylate kinase were detected after an extended exposure time ($>96\text{hrs}$), however the fainter crosslink species could not be resolved. Alternate detection by immunological detection of the poly-His tag using unlabelled adenylate kinase did not improve detection efficiency, even with the scaled up reactions.

These approaches failed to recover sufficient yield for detection by nickel-chelation. At this point attempts to further purify and isolate the crosslinked species were considered to be impractical, in part due to radioactive-decay of precursor. However this practical investigation did determine and confirm the characteristics of the hydrogenosome already observed by (Bradley 1997) with the *T. vaginalis* adenylate kinase preprotein. Additionally these data indicate the putative existence of discrete hydrogenosomal surface complexes, of which some mass determination has been made. These data are presented again the discussion, as well approaches to overcome the practical obstacles.

5 Bioinformatic analysis of the mitochondrial import system in *T. vaginalis*

The second approach to practical investigation in this study- the bottom up approach, provides the focus for the remaining results chapters in this project. This approach starts with an *in silico* investigation to potential candidates, which is the focus of this chapter.

To begin the overall structure of the translocase is re-introduced, with specific reference to the architectural features which will be sought for in the *in silico* investigation (5.1). The architecture of the system is then analysed complex by complex, first with the complexes of the outer membrane (5.2), then the inner (5.3) and finally the matrix (5.4). The specific elements of each complex are introduced in detail, and are followed by the bioinformatic analysis of their candidates in *T. vaginalis*. These investigations focus on providing data relating to the phylogeny and sequence features, where relevant relating to the definitive features of the complex. Finally this computational survey is summarised at the end of this chapter with a list of the collected candidates.

5.1 Elements of mitochondrial import model

The mitochondrial system is dependent on efficient import and targeting of its nuclear expressed preprotein products for function and viability. Within the mitochondrion there exist discrete membrane complexes which translocate preproteins across membrane compartments. This *in silico* investigation will focus on the four principle membrane translocases of the mitochondrion; TOM, SAM, TIM22, 23 complexes which transport the diverse number of nuclear encoded mitochondrial preproteins to their functional compartments (See Figure 5.1.1) as well as a matrix complex- PAM, which contributes to matrix import.

These membrane complexes share a number of common features, they are composed of translocases which conduct preproteins through membrane spanning pores, and additional proteins which function as receptors, assembly and regulatory factors, and proteins which recruit additional proteins to the translocon.

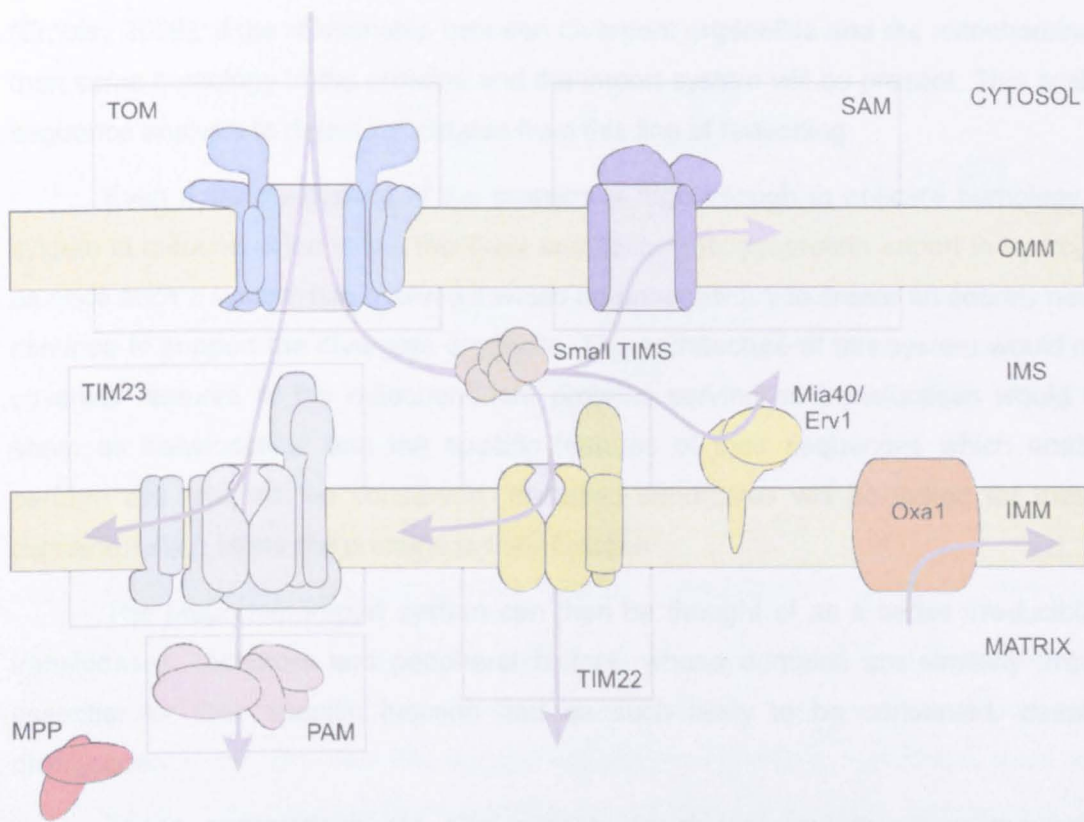


Figure 5.1.1 The principle translocons of the preprotein import machinery of the mitochondrion

The preprotein import system of the mitochondrion is composed of several discrete translocase complexes, each of which facilitates preprotein transport into a specific organellar compartment. These complexes are composed of the translocases with transport preproteins, but also receptors and other regulatory factors. In the mitochondrial model all proteins must traverse the TOM complex to enter the organelle. Outer membrane proteins are processed by the SAM complex, whereas proteins for the inner membrane and matrix are processed by the TIM22,23 complexes. Additional complexes such as PAM which assists in matrix import, and Oxa1, Erv1 complexes complete the functional aspects of this preprotein import system.

In *T. vaginalis* all hydrogenosomal proteins are nuclear encoded, it thus seems reasonable then to assume that the protein diversity of this organelle is supported by a preprotein import system which is able to transport proteins to specific hydrogenosomal compartments. The topological domains of the hydrogenosome are similar to the mitochondrion, and it is thus expected that the hydrogenosome needs a similar architecture of translocase complexes. Whilst the mechanistic nature of the import machinery is only partially characterized (Bradley 1997) we can assume that a functional system must be composed of translocases, receptors and other accessory factors as in mitochondria.

As discussed in the introduction this investigation will use the mitochondrial import system as a model to find *T.vaginalis* homologues. The relation of the hydrogenosome to the mitochondrion has been frequently discussed (Muller 1993; Bui & Johnson 1996; Tovar, Fischer & Clark 1999; Herrmann 2003) especially given the polyphyletic origins of these organelles

(Embley 2006). If the relationship between divergent organelles and the mitochondrion are valid then some homology to the proteins and the import system will be present. This section will use sequence analysis to detect candidates from this line of reasoning

Even if the divergence of the proteins is high enough to obscure homology, the import system in mitochondrion is still the likely architecture for preprotein import in hydrogenosomes, as once such a system has evolved it would be unnecessary to create an entirely new system to continue to support the divergent organelle. The architecture of this system would maintain the essential features of the mitochondrion, proteins serving as translocases would continue to serve as translocases, and the specific features of their sequences which enable them to perform this task will be conserved. Identified candidates will be tested for these essential domains, which relate the proteins to their function

The preprotein import system can then be thought of as a series irreducible elements, translocases, receptors and peripheral factors, whose domains are similarly irreducible and essential for their specific function and as such likely to be conserved, despite extreme divergences.

These assumptions are also suitable for the arguments where the origins of the hydrogenosome are independent or hybrid. As discussed in the introduction the ancestral population of proteins available for creation of the mitochondrial import model would also be available for secondary endosymbioses, and such systems might be expected to employ similar strategies, with the same ancestral genes, with the same domains. In the *in silico* analyses phylogenetic trees will be constructed to estimate the descent of *T. vaginalis* genes with respect to eukaryotic mitochondrial proteins, and their determined prokaryotic ancestors.

These rationalizations dictate a robust strategy to determining translocase candidates in the *T. vaginalis* hydrogenosome, whose analyses go beyond candidate selection and begin to examine the evolution of a complex molecular system.

5.2 Complexes of the outer membrane

The outer membrane of the mitochondrion has two protein complexes related to the targeting of mitochondrial preproteins, the TOM complex, and the SAM complex. Whilst structurally similar they occupy two distinct roles in preprotein trafficking.

TOM Complex

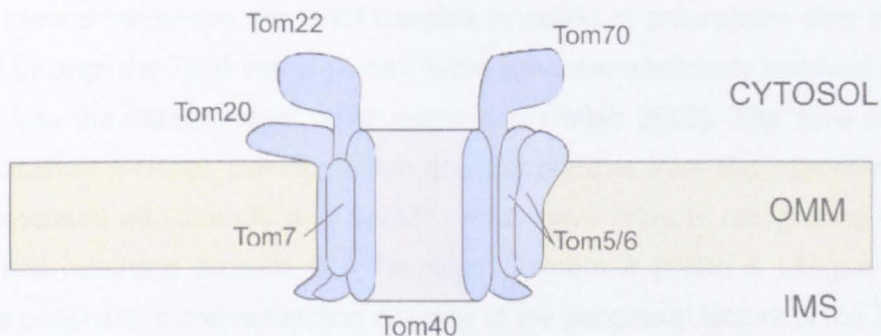


Figure 5.2.1 The TOM complex and its principle constituent proteins

All preproteins enter the mitochondrion through the TOM complex. It is composed of receptors (Tom20, 22, 70), and the translocase β -barrel channel protein Tom40. Preproteins are then conducted through the Tom40 pore in a partially unfolded state. In addition to the pore and receptors, additional proteins regulate the stability and assembly of the TOM complex (Tom5, 6, 7).

The TOM complex is the first complex any cytosolic preprotein must encounter to enter the organelle, it functions as the primary gateway for the mitochondrion (Ahting et al. 2001; Gabriel et al. 2003). This first translocase consists of a β -barrel membrane spanning protein, Tom40, which forms the membrane spanning pore (Ahting et al. 2001; Gabriel et al. 2003), this pore then associates with receptor proteins Tom20, 22, which recognise presequence motifs on preproteins, and together form the essential core of the TOM complex (Meisinger et al. 2001). This complex associates with other receptors, such as Tom 70 which assists in the translocation of presequence-less preproteins (Yamamoto et al. 2009), and the small proteins Tom5, 6, 7 which have roles in the assembly, stability and regulation of the TOM complex (Dembowski 2001; Schmitt 2005; Wideman et al. 2010).

SAM Complex

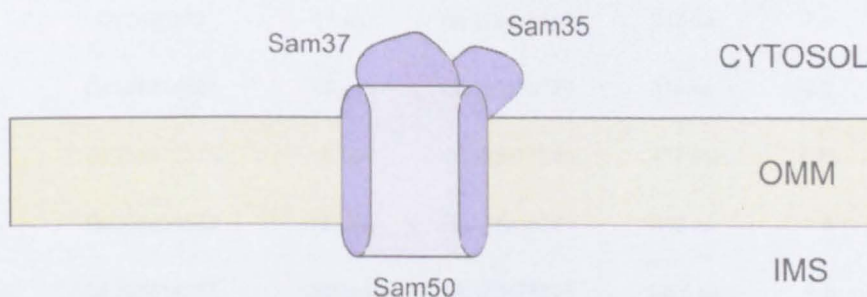


Figure 5.2.2 SAM complex and its protein components.

The SAM complex shares some common features with the TOM complex, the core translocase is the β -barrel protein Sam50, which accepts preproteins from the inter-membrane space, this complex is also assisted by Sam35,37 which

have roles in recognizing β -barrel signal sequences, and regulate the stability of the Sam50 allowing it to efficiently release proteins into the outer membrane.

The second translocon, the SAM complex is visited by preproteins after they have been translocated through the TOM complex, and is the essential machinery involved in the insertion of β -barrels into the mitochondrial outer membrane (Habib 2005). The core of this complex consists of another β -barrel, Sam50, which accepts proteins from the inter-membrane space, and that associates with Sam35 and Sam37 which have roles in recognising β -barrel signal sequences and releasing proteins into the outer membrane (Chan & Lithgow 2008). In this regard these peripheral proteins function similarly to the peripheral factors of the TOM complex.

5.2.1 Identifying *T. vaginalis* homologues to the outer membrane complexes

In this study the mitochondrial system of *S. cerevisiae* was used as the base for finding *T. vaginalis* protein homologues. Whilst other divergent bikonts are beginning to have the preprotein transport machinery characterized (for example *Entamoeba* (Dolezal et al. 2010)), the extensive characterization of the *S. cerevisiae* translocase system makes an attractive model system, especially with the extensive functional characterization data available for its components.

To identify if there are any highly conserved proteins in the outer membrane translocons a BLAST search was conducted using the *S. cerevisiae* proteins as a query against the *T. vaginalis* G3 genome. These results are shown below, the hits for the TOM complex are shown in Table 5.1 and those of the SAM complex in Table 5.2

| TOM component | <i>S. cerevisiae</i> query protein | Length | <i>T. vaginalis</i> protein | Length | Expect value (E) |
|---------------|------------------------------------|--------|-----------------------------|---------|------------------|
| Tom5 | GI:285815654 | 50 aa | no significant | | N/A |
| Tom6 | GI:1420173 | 61 aa | GI:123439157 | 647 aa | 7.4 |
| Tom7 | GI:285814581 | 60 aa | GI:123438120 | 214 aa | 2.3 |
| Tom 20 | GI:285812275 | 183 aa | GI:123498988 | 432 aa | 0.71 |
| Tom 22 | GI:285814523 | 152 aa | GI:154418221 | 855 aa | 1.9 |
| Tom40 | GI:285814207 | 387 aa | GI:123478859 | 1425 aa | 1.6 |
| Tom70 | GI:285814534 | 617 aa | GI:123437487 † | 579 aa | 7.00E-10 |
| | | | GI:154415632 | 345 aa | 2.00E-07 |

| | | |
|----------------|--------|----------|
| GI:123509584 | 264 aa | 6.00E-06 |
| GI:123412717 † | 561 aa | 2.00E-05 |
| GI:123976977 | 180 aa | 3.00E-05 |
| GI:123471216 † | 562 aa | 6.00E-05 |
| GI:123507524 † | 517 aa | 0.003 |
| GI:154414226 | 182 aa | 0.005 |
| GI:154419172 | 390 aa | 0.015 |
| GI:123457155 | 471 aa | 0.017 |
| GI:154418235 | 182 aa | 0.029 |
| GI:123448512 | 494 aa | 0.029 |
| GI:154412240 | 520 aa | 0.082 |

Table 5.1 BLAST results for TOM complex components

TOM complex components in *S. cerevisiae* were used to query the *T. vaginalis* genome using the p-BLAST algorithm, using the BLOSUM 45 matrix with a gap/extension penalty of 15:2. Results for each TOM complex component are shown where the expect value is <10, only the first sequence is shown for proteins where E≥1. Tom70 'homologues' identified in green are further investigated on p244

The BLAST search returned a great deal of hits to the *S. cerevisiae* preprotein receptor Tom70, these results were compelling given their weights (~65kDa) and resemblance to the unknown species determined in the import investigation. These results were critically analyzed to determine whether these results were genuine (see p244).

The multitude of results appear unusual, especially the quantity of short proteins. These results appear to arise because of multiple repeats of a TPR domain within Tom70, the BLAST search subsequently identified a number of proteins containing TPR motifs. Some intelligent discrimination can be applied to disregard the short *T. vaginalis* proteins, which leaves a small number of potential candidates. To further discriminate whether these results are genuine the domain organization of these results were compared against the *S. cerevisiae* query protein, and are shown in the appendices, p244, using the domain prediction tool CDD (Marchler-Bauer et al. 2011).

The domain architectures of the selected proteins are significantly different to the query protein which undermines their relation to the TOM receptor. However it has also been noted that Tom70 receptors have undergone a great deal of divergence (Suzuki et al. 2002) and whilst

this does not assist in the recognition of candidates, it might explain why no obvious candidates have been detected.

The lowest expectancy result for Tom40 is not likely to be a candidate translocase, given the great disparity in size, this would suggest that any TOM complex like translocases have diverged greatly in primary sequence similarity. The BLAST search did not return convincing candidates for the preprotein receptor proteins Tom20, 22, nor for the small assembly and regulation factors Tom5, 6, 7. Taken together the BLAST search using *S. cerevisiae* did not effectively find any convincing homologues for the TOM complex proteins.

| SAM component | <i>S. cerevisiae</i> query protein | Length | <i>T. vaginalis</i> protein | Length | Expect value (E) |
|---------------|---------------------------------------|--------|--------------------------------|---------|---------------------|
| Sam35 | GI:285809991 | 329 aa | GI:123468284 | 653 aa | 1.6 |
| Sam37 | GI:285814063 | 327 aa | GI:123424507 | 833 aa | 0.71 |
| Sam50 | GI:285814624 | 484 aa | GI:154414474 | 1241 aa | 0.82 |
| | | | GI:123456781 | 394 aa | 1.2 |

Table 5.2 BLAST results for the SAM complex

Using the *S. cerevisiae* model for the SAM complex, *S. cerevisiae* proteins were used to query the *T. vaginalis* genome using the BLAST-P algorithm for homologous sequences. The parameters for this search used the BLOSUM45 matrix and used a gap/extension penalty of 15:2. Sequences with an Expect value (E) <10 were collected and results are shown for sequences where $0 \leq E < 2$.

The BLAST results for the SAM complex showed similar results to those obtained from the search on the TOM complex. The candidate for the β -barrel protein Sam50 is much larger than expected, where as the cytosolic peripheral factors Sam35, 37 are very poor sequence matches and similarly much larger than expected.

The results of the BLAST search did not yield informative candidate choices, this does not infer that a mitochondrial system of outer membrane complexes does not exist, but rather that they have undergone extensive sequence divergence, or that the complex composition has changed. However the principles defined in the beginning of this section can guide further analysis of the outer membrane. We can assume that the outer membrane complexes of the hydrogenosome contains a similar kind of system architecture; that is, consisting of translocases which are β -barrels. We can use this principle to intelligently select outer membrane candidates.

In unpublished data by S.Dyall, mass spectrometry of the *T. vaginalis* hydrogenosomal proteome yielded a variety of unknown hydrogenosomal membrane proteins. On further analysis 11 of these proteins were predicted to have β -barrel secondary structure.

β -barrel proteins are a diverse group of proteins which originated from the original eukaryotic endosymbionts, and their localisation in eukaryotic organisms has remained restricted to the mitochondrion and other double membrane organelles (Schleiff & Soll 2005). The prokaryotic β -barrel Omp85 is thought to be the ancestor of the Sam50 proteins, whereas the nature of the Tom40 pore suggests a putative relation to porin type proteins in bacteria (Herrmann 2003; Schleiff & Soll 2005) (see Introduction p34). It seems reasonable that if the hydrogenosome, like the mitochondrion, is descended from α -proteobacterial or even any bacterial ancestor that β -barrels will still be the agents behind outer membrane translocation.

With this argument in mind 6 of the proteins identified in the mass spectrometry data (S. Dyall, unpublished) were further investigated. Five appear cladistically related, and belong to a family named in this project as the Hup3 family (Hydrogenosomal Unknown Protein). The sixth is an unusually long β -barrel candidate and seems unrelated to the Hup3 family which shall be called Hmp43 in this study (Hydrogenosomal Membrane Protein 43kDa). The essential features of these proteins are shown in Table 5.3

| Group | Protein | GI: number | Accession | Length |
|-------------|---------|--------------|--------------|--------|
| Hup3 family | Hup3a | GI:123474020 | XP_001320195 | 290 aa |
| | Hup3b | GI:123503140 | XP_001328448 | 297 aa |
| | Hup3c | GI:123387929 | XP_001299483 | 305 aa |
| | Hup3d | GI:123484034 | XP_001324169 | 308 aa |
| | Hup3e | GI:123433657 | XP_001308649 | 296 aa |
| Hmp43 | Hmp43 | GI:121914973 | EAY19766 | 398 aa |

Table 5.3 β -barrel proteins identified from hydrogenosomal membrane MS

This table shows the 6 β -barrel proteins examined in this investigation, the accession and GI numbers are shown, as well as the naming convention used in this thesis. 5 of these proteins belong to a shared group, the Hup3 family, and are characteristically of a similar size. Another protein, Hmp43, does not have any other closely related proteins in *T. vaginalis*, and is significantly larger than those in the Hup3 family.

These selected β -barrel proteins were tested using the Pred-TMBB (Bagos et al. 2004) tool, which uses a HMM algorithm trained with crystal structures from bacterial outer membrane proteins to determine likelihood of primary sequence submitted to being a β -barrel, and maps the predicted secondary structure of these sequences. All proteins except Hup3a fall within the threshold for being β -barrels by the Pred-TMBB tool. Hup3a narrowly missed this threshold

(2.965) by 0.023 (2.989), but its relation to the other Hup3 proteins warranted its inclusion into the set studied.

Whilst it is assumed that β -barrels will form the pore forming translocases in the outer membrane preprotein import complexes, β -barrels perform other roles in their endosymbiotically derived organelles. To affirm whether these candidate β -barrels are related to the preprotein import system a BLAST search was used to specifically score alignments between the hydrogenosomal β -barrel candidates and representative proteins of different eukaryotic β -barrel families. The results of this search are shown in Table 5.4

The results of this analysis are interesting, the hydrogenosomal proteins were aligned against preprotein import proteins from the mitochondrion (Tom40, Sam50), β -barrels which are not related to the TOM/SAM complexes (Por1, Mdm10), as well as chloroplast proteins (Toc75, OEP80), and *E. coli* Omp85. The Hup3 proteins show the greatest similarity to the Tom40 sequences, Hup3a, and Hup3e show especially strong sequence similarity to the *S. cerevisiae* Tom40, and even more pronounced relation to the recently identified *P. falciparum* Tom 40 protein (Mačasev et al. 2004), which is a closer common ancestor to *T. vaginalis* than *S. cerevisiae*. The Hup3 family has varying low similarity to other tested β -barrels, some intermittent relation to *S. cerevisiae* Por1 is seen but this might relate to the evolutionary history of Tom40 and Porin proteins (Herrmann 2003).

The sixth hydrogenosomal β -barrel candidate Hmp43 shows a different pattern of sequence scoring. Hmp43 has scores for Tom40 proteins which are generally weaker than for the Hup3 proteins, however it scores much more highly for *S. cerevisiae* Sam50, and *E. coli* Omp85 proteins, which are thought to be evolutionarily co-related (Schleiff & Soll 2005). These similarity patterns putatively suggest that these hydrogenosomal proteins have closest similarity to preprotein translocases in mitochondria, where the Hup3 family are closer in sequence similarity to the Tom40 β -barrels, and Hmp43 closer to Sam50 β -barrels.

| Hup3a | Hup3b | Hup3c | Hup3d | Hup3e | Hmp43 | GI: number | Protein | Organism |
|---------|---------|---------|---------|---------|---------|------------|---------|----------------------|
| 3E-06 | 1E-05 | 1E-01 | 4E-01 | 3E-09 | 2E-01 | 296004680 | Tom40 | <i>P. falciparum</i> |
| 5E-03 | 6E-01 | 3E-02 | 1E-02 | 2.0E-04 | 6.2E-01 | 6323859 | Tom40 | <i>S. cerevisiae</i> |
| 3E-01 | 4.6E-02 | 8.0E-01 | 6.7E-01 | 5.8E+00 | 4.0E-01 | 15232625 | Toc75 | <i>A. thaliana</i> |
| 4E-01 | 4.6E+00 | 1.0E+01 | 7.3E-02 | 7.5E+00 | 1.8E-01 | 6324302 | Sam50 | <i>S. cerevisiae</i> |
| 3.1E-01 | 4.5E-01 | 8.1E+00 | 6.2E-01 | 3.0E+00 | 3.3E-02 | 71164818 | Omp85 | <i>E. coli</i> |
| 7.6E+00 | 1.5E+00 | 8.0E-02 | 3.1E-01 | 2.1E+00 | 2.1E+00 | 18419973 | Oep80 | <i>A. thaliana</i> |
| 1.0E-03 | 1.3E+00 | 9.7E-01 | 2.8E-01 | 3.4E-01 | 4.0E+00 | 6324273 | Por1 | <i>S. cerevisiae</i> |
| 2.6E+00 | 2.3E+00 | 4.4E-01 | 4.2E-01 | 1.6E+00 | 2.6E+00 | 6319309 | Mdm10 | <i>S. cerevisiae</i> |

Table 5.4 BLAST scores of *T. vaginalis* β-barrel proteins against named β-barrel proteins from other organisms

T. vaginalis proteins were queried against membrane β-barrel proteins from other species. The BLAST search algorithm was executed with the BLOSUM45 matrix, with gap/extension penalty of 15:2. The alignment is scored as Expect (E) values, where strong alignments, based upon homology are reflected as low E values. Proteins which are related, and thus align well will typically have low E values. In this table the values have been colour coded, where colour intensity shows positive alignment.

5.2.2 Phylogeny of the *T. vaginalis* outer membrane protein homologues

Using BLAST pairwise alignment to score the *T. vaginalis* candidate β-barrel proteins hints at their possible homologues in *S. cerevisiae* and other mitochondrial systems. To build upon this analysis the β-barrel proteins were assembled together in a phylogenetic analysis. By combining the *T. vaginalis* β-barrel proteins together in one plot their relation to each other can be probed, as well as their place within other eukaryotic β-barrel families.

The analysis was constructed using BLAST-P searches against a non-redundant protein database. The algorithm parameters used a BLOSUM45 alignment matrix to better detect divergent proteins, with a gap extension penalty of 15:2, protein results with an expect value of <0.2 were included in the alignment. The *T. vaginalis* β-barrel proteins, Hup3a-e and Hmp43 were used as search queries as well as the *P. falciparum* Tom40 homologue, which scored particularly well against Hup3e. In addition representative members of the β-barrel families Omp85, Oep80, Toc75-III, Tom40, Sam50, Mdm10 and VDAC family were selected from NCBI's Homologene database.

The selected proteins and BLAST results were aligned in ClustalX, and neighbor-joining trees were calculated with 100 bootstraps. Species which returned multiple results were

reduced to single instances, and likewise not all of the proteins in the Homologene records for a given gene were included in the dendrogram (i.e. higher organisms). The final dendrogram includes the BLAST results for the *T. vaginalis* proteins as well as selected proteins from the *P. falciparum*, and can be seen in Figure 5.2.3.

The results of the phylogenetic analysis highlight some important properties of the *T. vaginalis* β -barrel proteins. The families described previously as Hup3 and Hmp43 are indeed parts of separate protein families, though both groups are close to the root of eukaryotic β -barrel family lineages. The results of the Hup3 BLAST search also recover two additional *T. vaginalis* proteins, these proteins are also predicted to have β -barrel properties, and perhaps reflect additional members of the Hup3 family as they clade within this group.

An additional analysis (Figure 5.2.4) using the BLAST algorithm to pairwise score Hup3 proteins against each other reveal that within the family there is some non-uniform divergence of primary sequence, perhaps reflecting that this family has arisen through repeated rounds of duplication and divergence. This phenomenon would explain the similarity of Hup3c, and Hup3d by a recent duplication event, and an older event where Hup3c, Hup3d diverged from Hup3a, it would also explain why there are many members of this β -barrel family. Whole genome duplication, and gene family expansion has already been described in *Trichomonas*, and thus it might seem reasonable that the same phenomenon is presented here (Carlton et al. 2007).

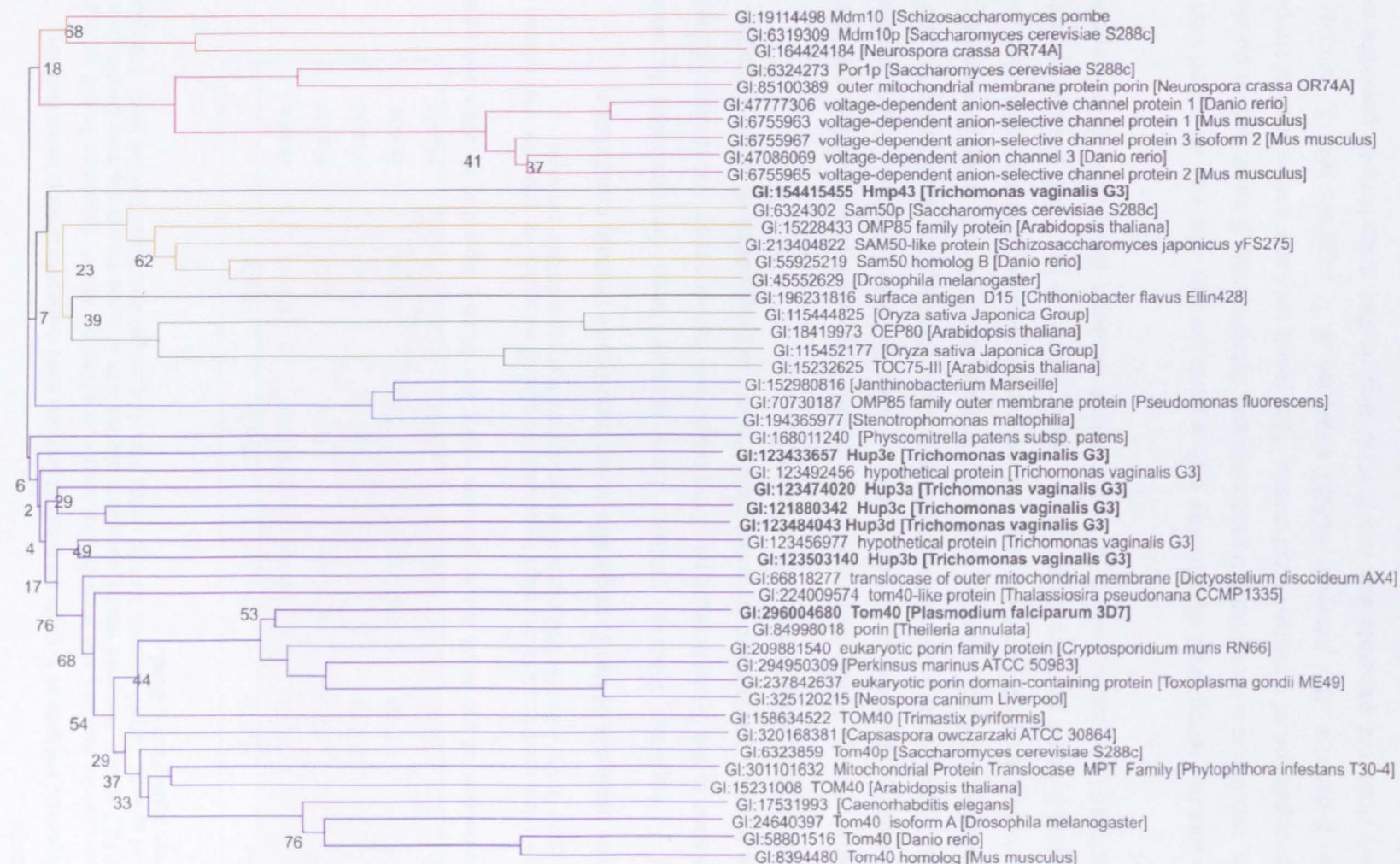
The β -barrel protein Hmp43 can be seen near the root of the Sam50 β -barrel family and close to the root of the Omp85 family. The evolutionary relation of these two families has already been posited (Schleiff & Soll 2005) whereby endosymbiont Omp85 family proteins were remodeled into the organellar Sam50 proteins (see Introduction p34). The position of the *T. vaginalis* Hmp43 protein, rooted deeply between these two clades, perhaps representing an, intermediate state between these families sharing some elements of both. If the hydrogenosome was modeled on a modern mitochondrion, with a modern Sam50 protein, its position within the Sam50 family might be more firmly rooted, whilst rapid divergence of Hmp43 would drive it to the periphery of the group, it does not necessarily drive it towards the Omp85 family.

The analysis indicates that the Hup3 family are most closely related to the Tom40 family of proteins. The BLAST search of the Hup3e protein returned several hits within *Plasmodium* identifying the putative Tom40 homologues (*P.falciparum* Tom40 E=0.005). The similarity to the *P. falciparum* sequence is encouraging evidence to support the position of the Hup3 family within the Tom40 β -barrel family, though their sequence similarity to the rest of the Tom40

group is highly divergent. Despite this the Hup3 family seem isolated from other β -barrel families, including the potentially evolutionary related Porin/VDAC family.

These results would suggest that the examined *T. vaginalis* β -barrel proteins are true eukaryotic β -barrels, and have some relation to specific eukaryotic β -barrel families. Their positions however do reflect their high divergence from other sequenced eukaryotes. These results explain the inability to find these proteins in the initial BLAST search.

1=0.01



↑ From previous page

Figure 5.2.3 Phylogenetic analysis of *T. vaginalis* β -barrel proteins with respect to other β -barrel families

T. vaginalis β -barrel proteins Hup3a-e and Hmp43 were used in combination with *P. falciparum* Tom40 as query sequences in a BLAST search. Returned results were then aligned with selected sequences from Omp85 (blue), Oep80 and Toc75-III (green), Tom40 (purple), Sam50 (gold), Mdm10 (orange) and VDAC families (red). The query sequences have been highlighted in **Bold**.

| Hup3a | Hup3b | Hup3c | Hup3d | Hup3e | |
|--------|--------|--------|--------|--------|-------|
| | 1.E-07 | 3.E-13 | 5.E-05 | 6.E-06 | Hup3a |
| 1.E-07 | | 1.E-04 | 9.E-02 | 1.E-05 | Hup3b |
| 3.E-13 | 1.E-04 | | 3.E-34 | 1.E-04 | Hup3c |
| 5.E-05 | 9.E-02 | 3.E-34 | | 2.E-04 | Hup3d |
| 6.E-06 | 1.E-05 | 1.E-04 | 2.E-04 | | Hup3e |

Figure 5.2.4 Inter-relation of the Hup3 family proteins using the BLAST algorithm to align and score sequences

The BLAST algorithm was used to score and align Hup3 proteins against each other to show the inter-relation of proteins within the Hup3 family. A BLOSUM62 matrix was used with default gap/extension penalties.

5.2.3 Sequence analysis of the *T. vaginalis* outer membrane protein candidates.

The candidate *T. vaginalis* β -barrel proteins have already been demonstrated to share phylogenetic relation to the β -barrel proteins of the preprotein translocases, the relation of these proteins can be further corroborated by examining the primary sequence of these *T. vaginalis* for motifs found in mitochondrial β -barrel translocases.

The domain library Pfam (Bateman et al. 2002), was used to screen the candidate sequences, the results generated by its search can be seen in Table 5.5 below. Pfam identified at least one significant domain for each candidate protein. For Hup3 proteins a Porin type 3 domain was identified for each member with a confidence value <0.019, strongly suggesting that all these proteins share a common and well conserved β -barrel domain associated with other mitochondrial β -barrel proteins. In addition a weak relation to the TonB β -barrel receptor family was identified for Hup3b, which is a membrane receptor β -barrel in gram-negative bacteria.

Hmp43 was not recognized significantly with a porin domain, but was identified with a mitofilin domain, mitofilin is a recognized transmembrane mitochondrial β -barrel with a putative role in structural regulation of christae mitochondria. This result confuses the relation of Hmp43 to a translocase β -barrel a little; however BLAST searches of *T. vaginalis* with *S. cerevisiae* mitofilin as query does not generate any hits E <3.0, and neither does Hmp43 generate any homologues to *S. cerevisiae* mitofilin in a reciprocal phylogenetic analysis.

| Protein | Domain | Pfam ID | E-Value |
|---------|-----------|------------|----------|
| Hup3a | Porin_3 | PF01459.15 | 6.10E-13 |
| Hup3b | TonB | PF03544.7 | 0.12 |
| | Porin_3 | PF01459.15 | 1.40E-16 |
| Hup3c | Porin_3 | PF01459.15 | 1.80E-07 |
| Hup3d | Porin_3 | PF01459.15 | 0.019 |
| Hup3e | Porin_3 | PF01459.15 | 5.80E-06 |
| Hmp43 | Mitofilin | PF09731.2 | 0.087 |

Table 5.5 A Pfam analysis of the β -barrel candidate protein sequences showing predicted domains.

The Pfam domain database was used to screen *T. vaginalis* outer membrane translocase candidates for evolutionarily conserved domains. Submitted sequences were then compared to consensus sequences within the Pfam database, and putative domain results scored. The Hup3 proteins return eukaryotic porin domains, which corroborate evidence of their β -barrel structure.

The β -barrel prediction program Pred-TMBB which was initially used to verify the *T. vaginalis* β -barrel proteins was further used to examine the secondary structure of the β -barrel proteins, to predict topology and transmembrane regions. An assemblage of the figures generated by the Pred-TMBB ANN can be seen in Appendix 7 p245- Hup3 family, Appendix 8 p246- Hmp43.

The Pred-TMBB program has some difficulty in dealing with sequences outside of the β -barrel structure, additional N or C-terminal domains outside of the membrane spanning segment maybe represented as additional membrane spans, this affect can be seen in Hup3a, c, e. These additional spurious spans also ultimately effect the determined topology of the protein.

On this basis making firm definitions on the insertion topology and β -barrel strand count cannot be made. However Hup3 proteins show a relatively similar topology and configuration, with tight inter-strand loops, but with considerable variation in specific loop length between proteins. In addition the C-terminal region of all Hup3, and Tom 40 proteins have a clearly related structure, terminating in two close transmembrane strands, Following these structures towards the N-terminal does suggest a generally conserved sequence of β -barrel strand pairs, though the additional length of the Tom40 proteins, and the variation in loop distance does obfuscate the structural relationships between the proteins shown.

Hmp43 has much more exaggerated inter-strand loops, a feature it shares in common with *S. cerevisiae* Sam50. Both β -barrel proteins share a C-terminal region with a group of tightly conserved β -strands, which is distinct from the doublets seen with Hup3/Tom40 C-

terminal region. Again tracing the transmembrane segments N-terminal of this region is difficult, an artifact perhaps of the posterior decoding method used in Pred-TMBB.

The topology predictions of Pred-TMBB partly corroborate the phylogenetic findings, where the Hup3 and Hmp43 proteins have distinct features, and that these features are unique to their corresponding groups (Tom40/Sam50). A common feature of the *T. vaginalis* proteins is that they are generally shorter than their proposed homologues. Whilst short N-terminal pre-sequences to hydrogenosomal soluble proteins are known (Bradley 1997) which would represent a reduced total sequence length compared to longer mitochondrial signal sequences, β -barrels are not known to have the same system of N-terminal sequences, and so sequence loss must be in the loss of terminal domains or reduction in inter-strand length.

Whilst eukaryotic β -barrel proteins do not have conserved N-terminal signaling sequences, they do contain conserved C-terminal motifs. These motifs are conserved across all eukaryotic β -barrel proteins and proteins of the SAM complex have been shown to interact with these sequences in their outer membrane protein insertion activities (Chan & Lithgow 2008).

This β -barrel signal motif was recognized in sequence alignments with other eukaryotic β -barrel protein families, and an alignment showing the conservation of these residues is shown in Figure 5.2.5. Whilst this feature alone cannot reconcile the *T. vaginalis* candidates to specific β -barrel families, they do indicate that this C-terminal sequence is conserved in *T. vaginalis*.

The findings of this section indicate that there are putative β -barrel proteins with homology present to mitochondrial outer membrane translocases within the *T. vaginalis* genome. However other aspects of the system are not easily found. The context of these results are revisited at the end of this section.

| GI: number | Species | Protein | Alignment | | | | | | | |
|--------------|---------|-----------|-----------|------------|----|-----|-----|------------------------------------|--|--|
| GI:6322077 | Sce | Por2 | PVH | | KF | GWS | LSF | SP | | |
| GI:6324273 | Sce | Por1 | PVH | | KL | GWS | LSF | DA | | |
| GI:296004680 | Pfa | Tom40 | IDTSG | KISVFTQDYS | | GFG | VSG | YIDYLNNDYKFGFMMHISPSQEQPQTAA | | |
| GI:6323859 | Sce | Tom40 | IDH | FKND T KI | | GCG | LQF | ETAGNQELLMLQQGLDADGNPLQALPQL | | |
| GI:123474020 | Tva | Hup3a | THLDH | KNAQ Y NF | | GLG | FQW | VENSTN | | |
| GI:123503140 | Tva | Hup3b | SFADH | FQKL Y SL | | GMA | VSV | RDTSST | | |
| GI:123387929 | Tva | Hup3c | SILDH | PAKN Y KL | | GLG | FYS | Q | | |
| GI:123484034 | Tva | Hup3d | GTLNH | KAKS Y TI | | GLG | FVF | NPDDFVST | | |
| GI:123433657 | Tva | Hup3e | CCLNH | LEAD Y SF | | GID | LSI | NQ | | |
| GI:121914973 | Tva | Hmp43 | SCDPS | WTSL A VA | | GVG | LTF | IQQQVKVEANLQKPIFQSGAKLNVLTYQLGITPA | | |
| GI:15232625 | Ath | Toc75-III | YAVDH | | NN | GTG | ALF | FRFGERY | | |
| GI:6324302 | Sce | Sam50 | HENDL | IRKG F QF | | GLG | LAF | L | | |
| GI:308153658 | Sce | Mdm10 | ENGNI | PVEP A KF | | GIQ | FQY | ST | | |
| GI:253775151 | Eco | Omp86 | AVSDI | RRSDFKTGT | | GVG | VRW | ESPVGPIKLDFAVPVADKDEHGLQFYIGLGPEL | | |

Figure 5.2.5 β -barrel signal sequences in *T. vaginalis* outer membrane β -barrels

Hmp43 and Hup3 proteins were aligned against selected β -barrel proteins from *S. cerevisiae* (Sce), *P. falciparum* (Pfa), *A. thaliana* (Ath) and *E. coli* (Eco). This Table shows the C-terminal region of the selected proteins, with spacing added to align the features of the C-terminal β -barrel signal. The signal is composed of a I/L-D/N-H motif (green), followed by a stretch of residues with one or more positive charges (red) and a G-X-G motif (orange). Hydrophobic residues (blue) are found immediately before and at positions +1,3 of the G-X-G motif. This motif is relatively well conserved across all β -barrels shown in this alignment, despite their diverse origins.

5.3 Complexes of the Inner Membrane

The inner mitochondrial membrane has two primary preprotein translocase complexes, both of which have functions in transporting preproteins into the matrix or the inner membrane. The degree of functional equivalence of these complexes varies according to the signal sequences present in individual preproteins, with some membrane proteins able to be processed by either complex (Yamano 2005) and others showing specificity for either TIM22 complex or TIM23 complex (Sirrenberg et al. 1996). Generally matrix soluble enzymes are processed by the TIM23 complex, whereas membrane insertion of inner membrane carrier proteins and other integral membrane proteins is mediated by the Tim22 complex.

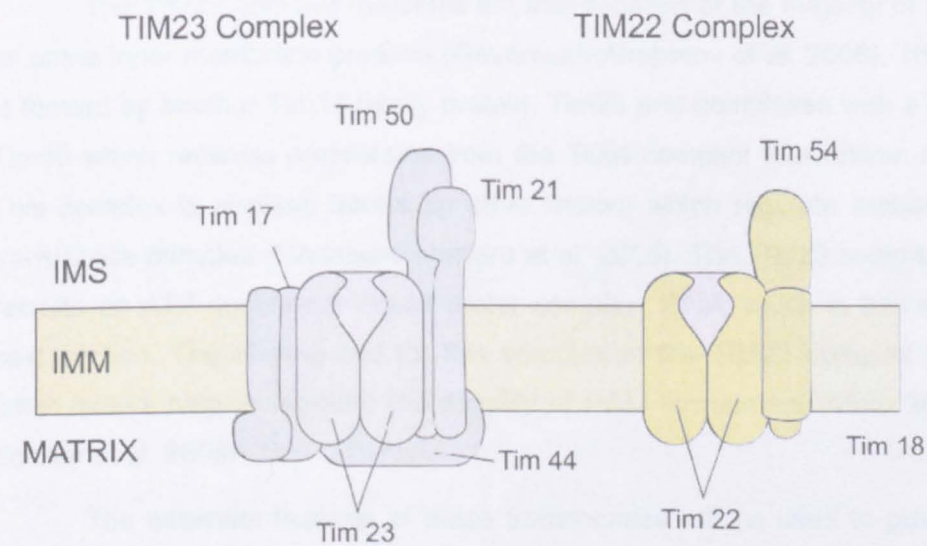


Figure 5.3.1 The mitochondrial inner membrane translocases complexes, TIM22 and TIM23

The inner mitochondrial membrane separates the inter-membrane space from the matrix. Preproteins in the inter membrane space which are either destined for inner membrane insertion, or for translocation into the organellar matrix are bound by chaperone proteins and directed to one of the two inner membrane translocase complexes TIM22, or TIM23.

Both translocons have some architectural similarity to the complexes in the outer membrane. Each consists of receptor proteins which associate with a core translocase, forming the essential core translocon, this structure is also bound by other protein factors which regulate the stability and assembly of the translocon. The inner membrane complexes are mechanistically different in the fact that they can employ a membrane potential to assist in preprotein import, as well as recruit additional matrix soluble complexes. The core translocase is also different in that it is not a β -barrel protein, but rather composed of proteins with a tim17

domain- containing 4 transmembrane α -helices. These proteins are theorized to be of eukaryote host origin (Introduction, p34).

The TIM22 complex consists of a core structure containing the Tim17-family translocase Tim22 which forms the transmembrane channel and accepts preproteins for membrane insertion (Sirrenberg et al. 1996; Peixoto et al. 2007) and Tim54 which although not directly involved with the preprotein import process contributes a scaffolding and structural role to the TIM22 complex (Hwang et al. 2007). This core complex peripherally associates with peripheral chaperone proteins in the inter-membrane space to accept preproteins, as well as additional integral membrane protein factor, Tim18 (Koehler et al. 2000).

The TIM23 complex mediates the translocation of the majority of matrix proteins as well as some inner membrane proteins (Gevorkyan-Airapetov et al. 2008). The translocase channel is formed by another Tim17-family protein, Tim23 and complexes with a larger receptor protein Tim50 which receives preproteins from the TOM complex (Gevorkyan-Airapetov et al. 2008). This complex is similarly bound by other factors which regulate stability and function of the translocase complex (Martinez-Caballero et al. 2006). The TIM23 complex is distinct that it also recruits an ATP-dependent import motor complex, PAM, which is further characterized in the next section. The binding site for this complex in the TIM23 complex protein Tim44, though other factors help to regulate the stability of PAM recruitment (Moro 2001; Hutu et al. 2008; Schiller et al. 2008).

The essential features of these translocases will be used to guide the investigation of homologues of these proteins in *T. vaginalis* in the following section.

5.3.1 Identifying *T. vaginalis* homologues to the inner membrane complexes

To identify *T. vaginalis* homologues from the *S. cerevisiae* mitochondrion a BLAST search was conducted. As before the BLOSUM45 matrix was used as a more sensitive detector of divergent proteins, the results of this search can be seen below, Table 5.6, contains the results for the TIM22 complex, and Table 5.7 the results for the TIM23 complex.

| TIM22 component | <i>S. cerevisiae</i> query protein | Length | <i>T. vaginalis</i> protein | Length | Expect value (E) |
|-----------------|------------------------------------|--------|-----------------------------|---------|------------------|
| Tim22 | GI:285810823 | 207 aa | GI:123477150 | 457 aa | 4.4 |
| Tim18 | GI:285813590 | 168 aa | GI:123451349 | 1597 aa | 0.71 |
| Tim54 | GI:285812846 | 478 aa | GI:123977093 | 372 aa | 0.35 |
| | | | GI:154414094 | 239 aa | 0.36 |
| | | | GI:154413683 | 140 aa | 0.89 |

Table 5.6 BLAST results for the TIM22 complex

Components of the TIM22 complex were sought in *T. vaginalis* using query proteins from the *S. cerevisiae* model. Proteins were identified using the BLAST algorithm using the BLOSUM45 matrix with a gap/extension penalty of 15:2. Results where $E \leq 1$ are shown, as well as the lowest scoring result where $E < 10$.

The results for the TIM22 complex do not identify any obvious homologues to the *S. cerevisiae* query proteins. Candidate proteins for Tim22 and Tim18 are greatly larger than expected, with poor BLAST scores. Multiple results were obtained for Tim54- however the result proteins are small, seemingly co-related by a poorly defined super-family domain.

The results for the TIM23 complex are a little more encouraging than the outer membrane data. Results for the Tim50 protein are unexpectedly numerous, though all results can be discounted on the basis of length, and that the results originate from an common NLI domain, similar to the effect observed with *S. cerevisiae* Tom70.

| TIM23 component | <i>S. cerevisiae</i> query protein | Length | <i>T. vaginalis</i> protein | Length | Expect value (E) |
|-----------------|------------------------------------|--------|-----------------------------|--------|------------------|
| Tim23 | GI:285814664 | 222 aa | GI:154418731 (Phata) | 160 aa | 0.074 |
| | | | GI:123445974 | 1098 a | 0.14 |
| | | | GI:123473678 | 548 aa | 0.23 |
| Tim17 | GI:285812760 | 158 aa | GI:123431948† | 152 aa | 0.25 |
| | | | GI:123389952 | 366 aa | 0.77 |
| Tim21 | GI:285812230 | 239 aa | GI:123471617 | 226 aa | 0.75 |
| Tim50 | GI:285815476 | 476 aa | GI:123402964 | 287 aa | 1.00E-24 |
| | | | GI:154411691 | 322 aa | 3.00E-23 |
| | | | GI:123457873 | 323 aa | 2.00E-22 |
| | | | GI:123496080 | 288 aa | 5.00E-20 |
| | | | GI:123434330 | 324 aa | 2.00E-19 |
| | | | GI:123483322 | 210 aa | 1.00E-18 |
| | | | GI:123479490 | 315 aa | 6.00E-18 |
| | | | GI:123497759 | 210 aa | 2.00E-16 |
| | | | GI:123433208 | 207 aa | 8.00E-15 |
| | | | GI:123410353 | 244 aa | 9.00E-15 |
| | | | GI:123430726 | 223 aa | 1.00E-14 |
| | | | GI:123384365 | 242 aa | 3.00E-14 |
| | | | GI:123506276 | 188 aa | 4.00E-13 |
| | | | GI:123454430 | 218 aa | 2.00E-11 |
| | | | GI:123428297 | 244 aa | 3.00E-11 |
| | | | GI:123404051 | 205 aa | 1.00E-10 |
| | | | GI:154413895 | 198 aa | 3.00E-07 |
| | | | GI:123478332 | 192 aa | 4.00E-06 |
| | | | GI:123496394 | 397 aa | 0.007 |

Table 5.7 BLAST results for the TIM23 Complex

The BLAST algorithm was used to search for *T. vaginalis* proteins homologous to the TIM23 complex in *S. cerevisiae*. The search was run using the BLOSUM45 matrix and a gap/extension penalty of 15:2. Returned sequences with $E < 1$ are shown. Tim50 generated a large number of results from a conserved NLI family domain feature. Tim44 is not shown in this table- but is shown in table along with other proteins of the PAM complex.

S.cerevisiae Tim21 also seems to detect a *T. vaginalis* protein of approximately the same size, however the two proteins show very different domain composition. Similarly although *S. cerevisiae* Tim17 appeared to give a convincing homologue with an appropriate size, but further bioinformatics analysis using a previously identified *Trypanosoma brucei* Tim17 (Singha et al. 2008) did not confirm this identification, and domain structures between query and result protein did not conserve the Tim-superfamily domain.

However the BLAST search identified a very convincing homologue to *S. cerevisiae* Tim23, which has been termed here as 'Phata' (Putative Hydrogenosomal Transporter-A). Whilst the protein in the BLAST result is rather shorter than the query protein, it has a clear Tim17 domain, a definitive feature of the mitochondrial inner membrane translocases.

The existence of a single convincing homologue prompted a search for other homologous proteins in *T. vaginalis* given that Tim23 is one of a family of conserved proteins whose members participate in more than one inner membrane complex.

Given that the Tim17-superfamily domain is the essential feature of these translocases, the domain database Pfam was used to screen for other *T. vaginalis* proteins containing Tim17 domains. The Pfam database identified a second Tim-17 domain containing *T. vaginalis* protein, which has been included in this investigation as 'Phatc'.

Finally using these two proteins as queries in a BLAST search (using a BLOSUM45 matrix) a third *T. vaginalis* protein was identified with some homology to both Phata,c, and is included in this study as 'Phatb'. The essential features of these candidates are shown in Table 5.8, and are shown in relation to *S. cerevisiae* Tim-17 family membrane translocases. Given that the Tim17 domain is the definitive feature of this type of translocase, Pfam was further used to score domains within the *T. vaginalis* Phat proteins, and the results can be seen in Table 5.9.

| Protein | Species | Length | GI: number |
|---------|----------------------|--------|--------------|
| Tim17 | <i>S. cerevisiae</i> | 158 aa | GI:6322318 |
| Tim22 | <i>S. cerevisiae</i> | 207 aa | GI:6319984 |
| Tim23 | <i>S. cerevisiae</i> | 222 aa | GI:6324344 |
| Phata | <i>T. vaginalis</i> | 160 aa | GI:121916618 |
| Phatb | <i>T. vaginalis</i> | 159 aa | GI:121890196 |
| Phatc | <i>T. vaginalis</i> | 139 aa | GI:121906375 |

Table 5.8 Essential features of the Phat family of proteins

In addition to Phata two other inner membrane translocases were identified, Phatb and Phatc, the candidates essential properties are shown above. The properties of the TIM17 super-family homologues (Tim17,22,23) of *S. cerevisiae* are also shown for comparison. The *T. vaginalis* Phat proteins are typically shorter, similar in size to *S. cerevisiae* Tim17, with less variation in size within the group

| Protein | Domain | Pfam ID | E-Value |
|---------|-----------------|------------|----------|
| Phata | Pfam-B_3942 | PB003942 | 5.60E-07 |
| | Tim17 | PF02466.12 | 1.50E-18 |
| Phatb | Tim17 | PF02466.12 | 7.20E-06 |
| Phatc | Tim17 | PF02466.12 | 2.30E-14 |
| | MerE | PF05052.5 | 0.0037 |
| | Rick_17kDa_Anti | PF05433.8 | 0.0025 |

Table 5.9 Predicted domains of the Phat family of proteins by Pfam

This table lists the statistically significant ($E < 1.0$) domains predicted for the Phat family of proteins by the protein domain search engine Pfam. The Pfam E values represent a confidence value produced by the HMM method of aligning the query sequence against a consensus for that domain. These consensus sequences are built from alignments, whose Pfam IDs are shown.

The Pfam search did positively rate all Phat proteins with the essential Tim17-superfamily domain to at least $< 1 \times 10^{-5}$, and in all Phat proteins the Tim17 domain was the most confident result. Pfam also identified some additional domains within the Phat proteins, some of these are autonomously generated domains which are currently unlabelled with any specific relation (Pfam-B_3942), however two weaker results were identified for Phatc, identifying a MerE, 'Rick_17kDa_Anti' domain both of which are prokaryotic domains, however neither of

these has a strong relation to protein translocation activity, nor are they extensively characterized, however both contain transmembrane α -helices.

The core Tim17 domain generates a membrane spanning channel through the close interaction of four transmembrane α -helices, these helices have already been characterized in *S. cerevisiae* inner membrane translocases, and their existence was posited for the *T. brucei* Tim17 homologue by scanning the primary sequence for regions of hydrophobicity (Singha et al. 2008). This same method was applied to screen the *T. vaginalis* Phat proteins using the TopPred program, using a Kyte and Doolittle hydrophobicity scale (Kyte & Doolittle 1982) and a 5aa scan window, the results for the Phat protein proteins can be seen in Figure 5.3.2, and a comparative set of results using the same program settings for the *S. cerevisiae* Tim17, 22, 23 can be seen in Appendix 10 p248.

The predicted hydrophobicity profile for the Phat proteins putatively suggests the existence of membrane spanning sections, with a comparable profile to the *S. cerevisiae* proteins. All the proteins from *S. cerevisiae* and *T. vaginalis* share a similar C-terminal region with four predicted membrane spanning regions, corresponding to the position of the predicted Tim17 domain. Phata and *S. cerevisiae* Tim22 are predicted an additional membrane span in their N-terminal regions however this is not known to associate with an additional transmembrane helix. The similarity in profiles might suggest some degree of relation between specific Phat and *S. cerevisiae* translocase proteins.

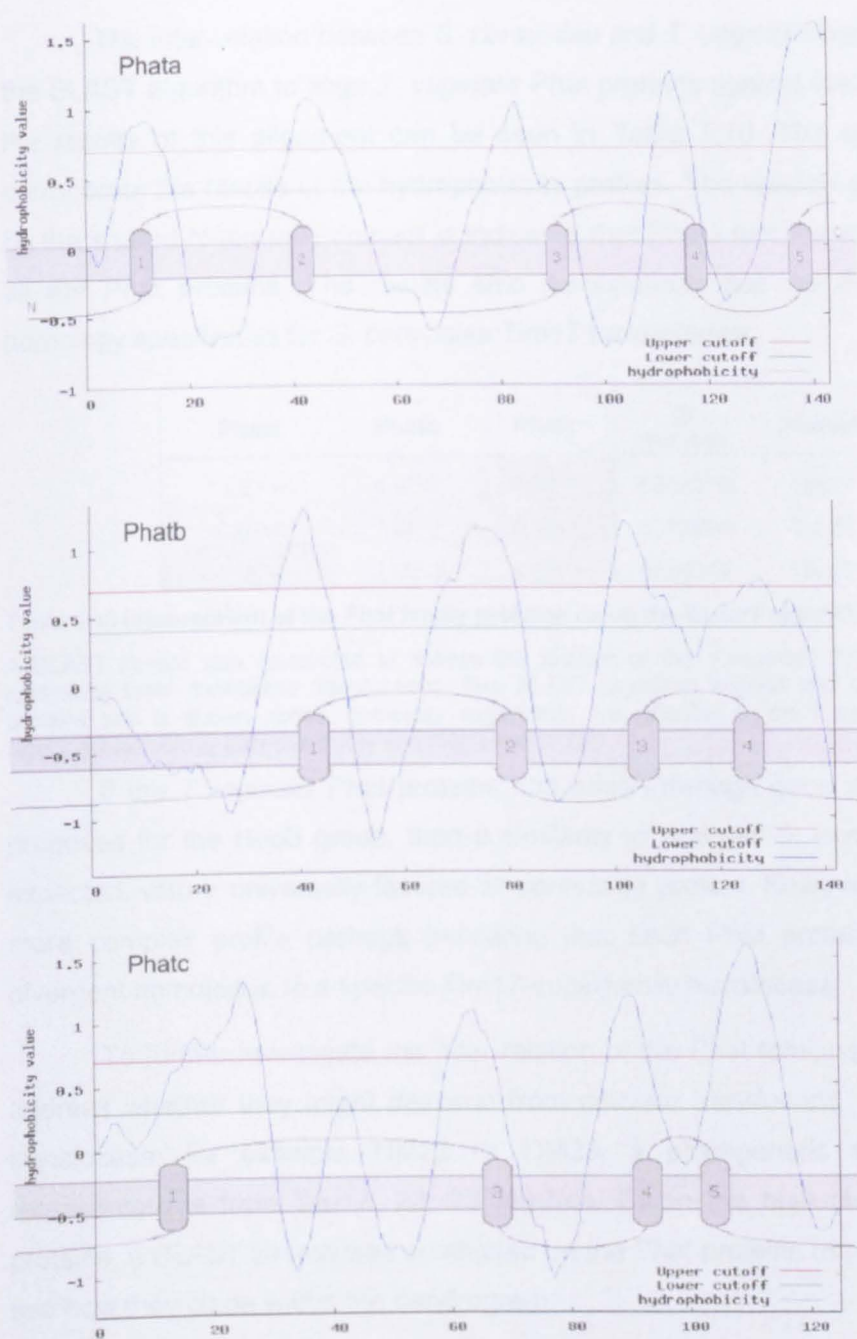


Figure 5.3.2 TopPred prediction of the transmembrane regions of the Tim17 domain in the Phat family proteins

The transmembrane region prediction program TopPred was used to predict transmembrane α -helices of the Tim17 domain within the Phat proteins. Putative (green) and certain cutoffs (red) were used with values of 0.3 and 0.7. A window size of 5 was used to detect the short loops present within this tightly packed domain, and the Kyte and Doolittle hydrophobicity scale was used to rate the hydrophobicity of the primary sequence (Kyte & Doolittle 1982). The Tim17 domain contains 4 α -helices which are clearly detected in each of the Phat proteins, an additional transition is detected in Phata, this effect is also seen in *S. cerevisiae* Tim 22 (see Appendix 10 p248) Which is likely to arise from an N-terminal region before the Tim17 domain, as Tim22 is known to have only 4 TM helices.

The inter-relation between *S. cerevisiae* and *T. vaginalis* translocases was probed using the BLAST algorithm to align *T. vaginalis* Phat proteins against their *S. cerevisiae* counterparts, the results of this alignment can be seen in Table 5.10. The results of BLAST search do corroborate the results of the hydrophobicity profiles. The relation of Phata to Tim23 suggested by the shared N-terminal domain is indicated that Phata has the strongest relation to Tim23 of all the Phat proteins. The results also demonstrate that the Phat proteins show different homology specificities for *S. cerevisiae* Tim17 translocases.

| Phata | Phatb | Phatc | gi number | Protein | Organism |
|---------|--------|--------|-----------|---------|----------------------|
| 1.0E+1< | 4.8E-2 | 3.5E-2 | 6322318 | Tim17 | <i>S. cerevisiae</i> |
| 2.8E+0 | 7.9E-1 | 6.7E-1 | 6319984 | Tim22 | <i>S. cerevisiae</i> |
| 4.0E-6 | 2.0E-3 | 1.0E-3 | 6324344 | Tim23 | <i>S. cerevisiae</i> |

Table 5.10 Inter-relation of the Phat family proteins using the BLAST algorithm to align and score sequences
A BLAST search was conducted to assess the relation of the *T.vaginalis* Tim17 containing proteins to the *S. cerevisiae* inner membrane translocases. The BLAST algorithm aligned and scored the homology between the proteins and is shown above, homology alignments are reflected in the E-value, where a low value suggests significant homology between query and reference protein.

If the *T.vaginalis* Phat proteins had arisen through gene duplication, similar to what is proposed for the Hup3 group, then a similarity to a single *S. cerevisiae* translocase might be expected, with a universally favored *S. cerevisiae* protein. However the Phat proteins exhibit a more complex profile perhaps indicating that each Phat protein might represent a greatly divergent homologue to a specific Tim17-superfamily translocase.

To further investigate the inter-relation of the Phat family of proteins, and to specifically address whether they might descend from discrete translocons, or instead relate to a single translocase, for example TIM22 or TIM23, a phylogenetic tree was constructed using representatives from Tim17, 22, 23 families. Given the high divergence of the *T. vaginalis* proteins, a BLAST search was conducted on the Phat proteins to generate related sequences to see how they clade within the dendrogram.

The results suggest that there is some divergence within the Phat proteins. Phata,c seem related, and clade deep within the TIM23 family. Phatb seems to relate more closely to the Tim17 family than to the other two Phat proteins.

Whilst the deeply divergent Phat proteins are too divergent to relate to the mitochondrial system to guess their equivalent translocase complex, it would appear that Phata, c would be likely to share a similar function, perhaps as Tim22, 23 do, and Phatb would occupy a different functional role as Tim17 does within the mitochondrial model.

5.4 The PAM motor and accessory proteins to the preprotein translocase system

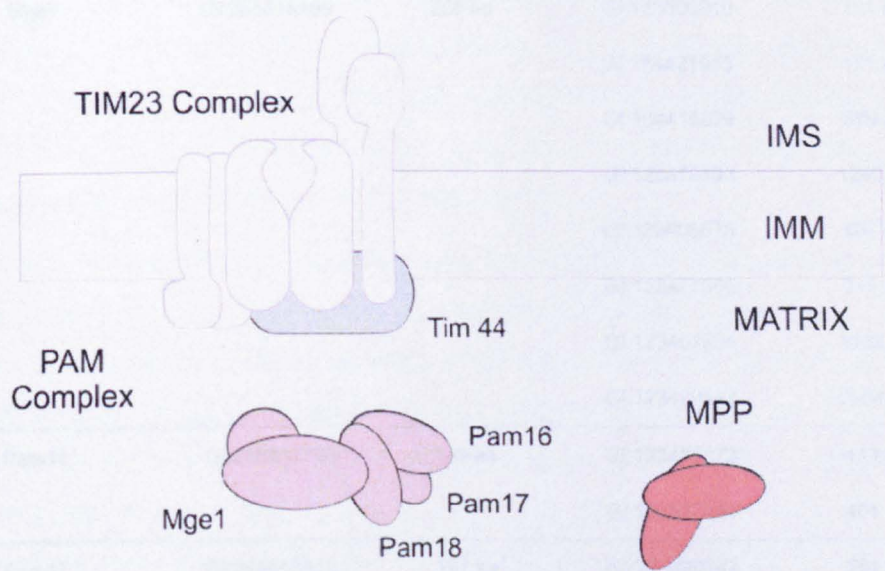


Figure 5.4.1 The structural relation of the PAM motor, to its binding site Tim44 on the TIM23 complex

Complete preprotein import through the TIM23 complex depends on the presence of membrane potential and ATP. The ATP dependence of this complex is generated from its recruitment of the PAM motor complex which uses ATP and mtHsp70 to fully translocate proteins into the organellar matrix. The PAM motor is composed of multiple small proteins with chaperone activity, and interacts with the TIM23 through numerous docking sites, particularly Tim44. Fully translocated preproteins are then matured with a presequence protease- MPP.

Complete translocation of matrix destined preproteins through the TIM23 requires the assistance of an accessory motor complex PAM (presequence translocase-associated motor), (Jensen & Johnson 2001; Truscott et al. 2003). Whilst the proteins of the PAM complex do not form a translocase in their own right, their role is essential for the function of the TIM23 complex, and confers the ATP requirement for preprotein import, these requirements appear to be also necessary in *T.vaginalis* hydrogenosomes (Bradley 1997). With potential candidates for the Tim17-superfamily translocases, a bioinformatic analysis of the PAM motor was conducted.

| PAM component | <i>S. cerevisiae</i> query protein | Length | <i>T. vaginalis</i> protein | Length | Expect value (E) |
|---------------|------------------------------------|--------|-----------------------------|---------|------------------|
| Mge1 | GI:285815109 | 228 aa | GI:123506910 | 191 aa | 3.00E-10 |
| | | | GI:154421943 | 191 aa | 3.00E-10 |
| | | | GI:154416829 | 829 aa | 0.19 |
| | | | GI:123478193 | 1240 aa | 0.21 |
| | | | GI:123406678 | 467 aa | 0.28 |
| | | | GI:123477366 | 215 aa | 0.38 |
| | | | GI:123407834 | 1165 aa | 0.48 |
| | | | GI:123409642 | 2354 aa | 0.94 |
| Pam16 | GI:285812798 | 149 aa | GI:123437973 | 111 aa | 0.025 |
| | | | GI:123507401 | 401 aa | 0.93 |
| Pam17 | GI:285813319 | 197 aa | GI:123429247 | 751 aa | 0.024 |
| | | | GI:123440199 | 756 aa | 0.032 |
| | | | GI:123252691 | 358 aa | 0.44 |
| | | | GI:123244006 | 344 aa | 0.65 |
| Pam18 | GI:285813430 | 168 aa | GI:154417601 | 117 aa | 2.00E-13 |
| | | | GI:123435417 | 191 aa | 0.028 |
| | | | GI:123437973 | 111 aa | 0.037 |
| | | | GI:123472833 | 335 aa | 0.23 |
| | | | GI:123405686 | 226 aa | 0.45 |
| Tim44 | GI:285812624 | 431 aa | GI:123446418 | 326 aa | 7.00E-04 |
| | | | GI:154419826 | 697 aa | 0.11 |
| | | | GI:123503850 | 390 aa | 0.16 |
| | | | GI:123974654 | 304 aa | 0.8 |

Table 5.11 BLAST results for the PAM Complex

Candidates for the PAM motor complex were identified using the BLAST algorithm using the BLOSUM45 matrix and gap/extension penalty of 15:2. Returned sequences where the expect value <1.0 are shown above. All *S. cerevisiae* queries return high scoring sequences in *T. vaginalis*, mtHsp70 homologues were also found, but not shown here due to the great number of significant returned results (~90), these are instead shown separately Appendices (p249 p250 p251). These data would suggest that the PAM motor elements are more highly conserved than other translocon complexes examined here, possibly deriving from the strongly conserved chaperone role that many of these proteins participate in.

The PAM motor is a soluble protein complex which is recruited to the translocon through interactions with the matrix faces of the TIM23 complex proteins, particularly Tim44, and Tim23 (Hutu et al. 2008) and interacts principally with PAM proteins, Pam16, 17, 18 (Frazier et al. 2004; van der Laan et al. 2005; Schiller 2009). The translocase activity imparted by the PAM complex comes through interaction with mtHsp70 with the emerging preprotein chain, this interaction is assisted by Mge1 which acts as a nucleotide exchange factor for mtHsp70 (Sakuragi et al. 1999; Nowack & Melkonian 2010).

Preproteins which have been fully translocated into the matrix often undergo signal peptide cleavage which is performed by a matrix processing peptidase (MPP), this protein is responsible for the mass reduction seen in mitochondrial matrix proteins, and for the *T. vaginalis* hydrogenosomal proteins (Brown et al. 2007). Whilst this protein has already been investigated its relation to the PAM will be briefly investigated here. Similarly the existence of a *T. vaginalis* mtHsp70 has already been ascertained in the context of proving the secondary loss of mitochondria from deep branching eukaryotes (Germot et al. 1996; Emelyanov & Goldberg 2011), but is further investigated here in its PAM context.

A bioinformatic analysis of the PAM components was made for *T. vaginalis* using the same approach as for other parts of the translocase system using the BLAST algorithm. The results of this search can be seen in Table 5.11. Results for mtHsp70 and MPP are analyzed separately, as these have previously been identified in *T. vaginalis*.

The results of the BLAST search are surprising, and very different to the searches performed on the other complexes of the import model. For each element with the exception of Pam17, there appears to be at least one significant homologue present in *T. vaginalis*. These data would also reconcile with the existence of a mtHsp70 in the hydrogenosome (Germot 1996) and support a model whereby the whole of the PAM machinery and its peripheral proteins including the MPP are conserved within the hydrogenosome. This also corroborates the similarities for import between mitochondria and hydrogenosomes.

Whilst a homologue to mtHsp70 has already been identified (Germot et al. 1996) in *T. vaginalis* an investigation was performed to examine the diversity of the DnaK related proteins, which have various roles in chaperone activity. In this analysis DnaK related proteins from *S. cerevisiae* were used to identify similar sequences in *T. vaginalis* using a BLAST search. The results of the search reveal a very large expansion of these related proteins in *T. vaginalis* possibly as a result of repeated rounds of genome duplication.

A group of proteins clading with *S. cerevisiae* mtHsp70 and BIP indicate that candidates do exist for these two different molecular chaperones which have functions in ER and mitochondria. The mtHsp70 homologues including the protein previously identified as the *T. vaginalis* mtHsp70 homologue were further used in a phylogenetic analysis which shows the suggested position of *T. vaginalis* mtHsp70, and the protein's relation to α -proteobacteria, reproducing previous findings (Germot et al. 1996).

Of similar interest to the highly conserved mtHsp70 *T. vaginalis* has a highly conserved β -subunit of the mitochondrial processing peptidase, which would suggest an presequence protease system at work in the hydrogenosome, which is consistent with experimental work. The divergence of the MPP β -subunit is sufficient to make a meaningful phylogenetic analysis from the divergence of the mitochondrial MPP β -subunit.

0.01

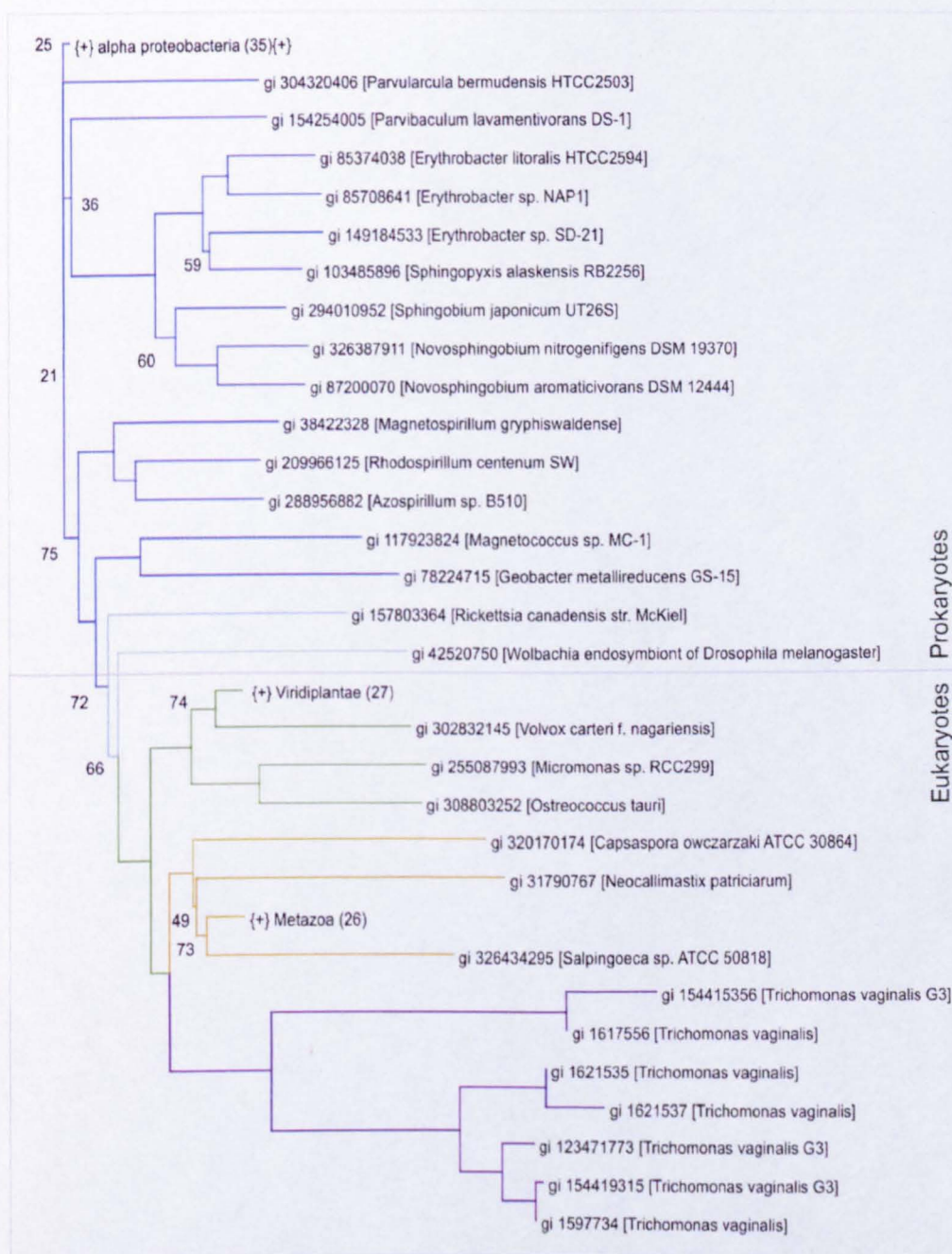


Figure 5.4.2 Phylogenetic analysis of the identified mtHsp70 candidates

Candidates for the hydrogenosomal mtHsp70 derived from previous phylogenetic analysis (See Appendix 12 p250, Appendix 13 p251) were used to construct a phylogenetic tree. The BLAST algorithm was used to identify homologous proteins to the *T. vaginalis* candidates, and because of the high degree of sequence conservation of the mtHsp70 sequences the BLOSUM80 matrix was used. The dendrogram has been coloured to differentiate between prokaryotes (blue), and plant (green), and metazoan (orange) clades. *T. vaginalis* proteins are shown in purple, and are seen as a deep branching group within the eukaryotic clade.

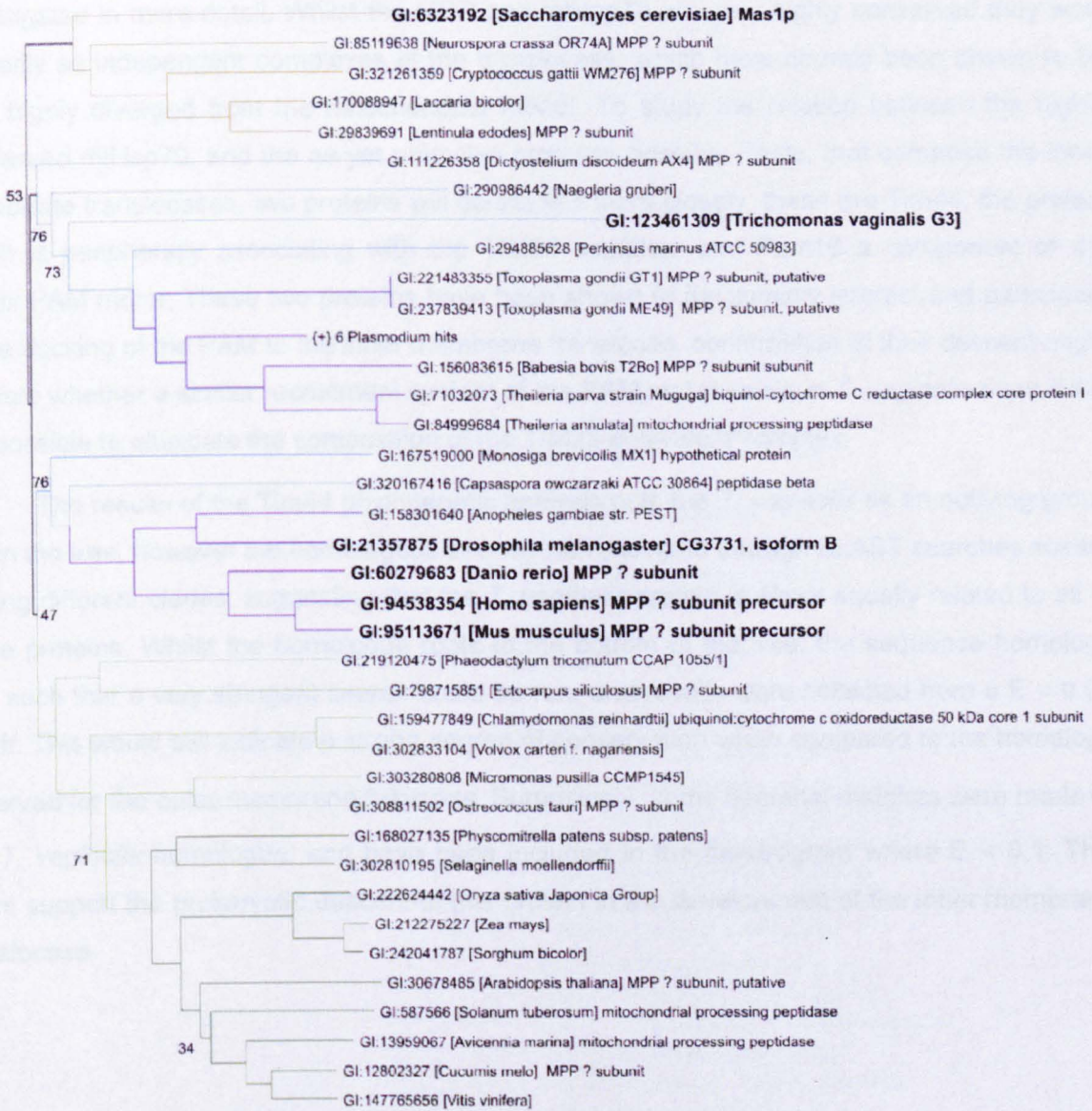


Figure 5.4.3 A phylogenetic analysis of the *T. vaginalis* β -MPP subunit.

A BLAST search was performed on the homologue to the β -MPP subunit found in *T. vaginalis* (black, bold) using a BLOSUM80 matrix with default gap penalties. The first 50 results (all <E-30, black) were then assembled in an alignment with the *T. vaginalis* homologue in addition to some known β -MPP subunits from other species (Bold). After the alignment a phylogenetic tree was constructed using the Neighbour-Joining algorithm, which was then bootstrapped 100 times.

The results of the phylogenetic analysis indicate that the *T. vaginalis* MPP clades from other extant organisms which possess mitochondria. This would suggest that the MPP subunit has diverged from a mitochondrial protein, and not that it shares a common phylogeny with an ancestral organellar endosymbiont.

In addition to these highly conserved proteins, two more divergent proteins were investigated in more detail. Whilst the MPP and mtHsp70 are very highly conserved they work primarily as independent complexes of the translocase, which have already been shown to be very highly diverged from the mitochondrial model. To study the relation between the highly conserved mtHsp70, and the as yet unknown proteins- possibly Phats, that compose the inner membrane translocases, two proteins will be studied more closely, these are Tim44, the protein which is peripherally associating with the TIM23 complex, and Pam16 a component of the matrix PAM motor. These two proteins have been shown to functionally interact and participate in the docking of the PAM to the inner membrane translocon, confirmation of their descent might indicate whether a similar recruitment system of the PAM motor exists in *T. vaginalis* even if it is not possible to elucidate the composition of the TIM23 equivalent complex.

The results of the Tim44 phylogenetic analysis puts the *T. vaginalis* as an outlying group within the tree, however the homologous sequences recovered through BLAST searches scatter among different clades, suggesting that the *T. vaginalis* protein is about equally related to all of these proteins. Whilst the homologue roots to the bottom of the tree, the sequence homology was such that a very stringent search could be run, and results were collected from a $E < 0.01$ cutoff. This would still indicate a strong degree of conservation when compared to the homology observed for the outer membrane β -barrels. Surprisingly, some bacterial matches were made to the *T. vaginalis* homologue, and have been included in the dendrogram where $E < 0.1$. This might support the prokaryotic descent of this protein in the development of the inner membrane translocase.

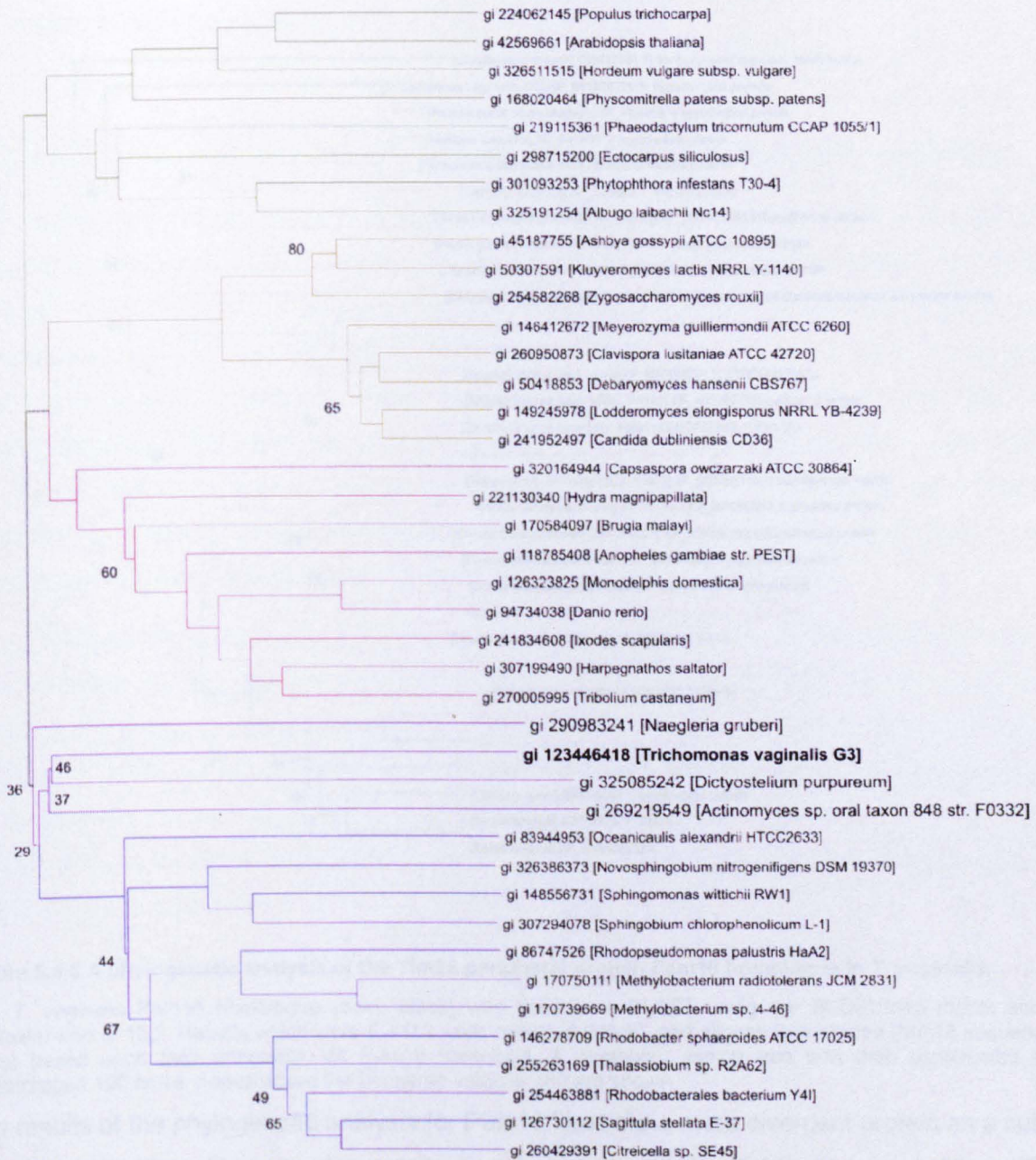


Figure 5.4.4 A phylogenetic analysis of the Tim23 peripheral protein Tim44 homologue in *T. vaginalis*.

The phylogeny of the *T. vaginalis* homologue to the Tim44 protein (black, bold) was investigated by constructing an alignment from homologous sequences collected from BLAST search of the *T. vaginalis* protein. Because of the high level of conservation of the proteins the BLOSUM80 matrix was used with default gap extension penalties. An alignment was constructed in ClustalW and a Neighbour-Joining tree calculated with 100 bootstraps. Branches with a bootstrap value of <80 are shown at their nodes. Clades have been colour coded, plantae, fungi, animalia, bacteria.

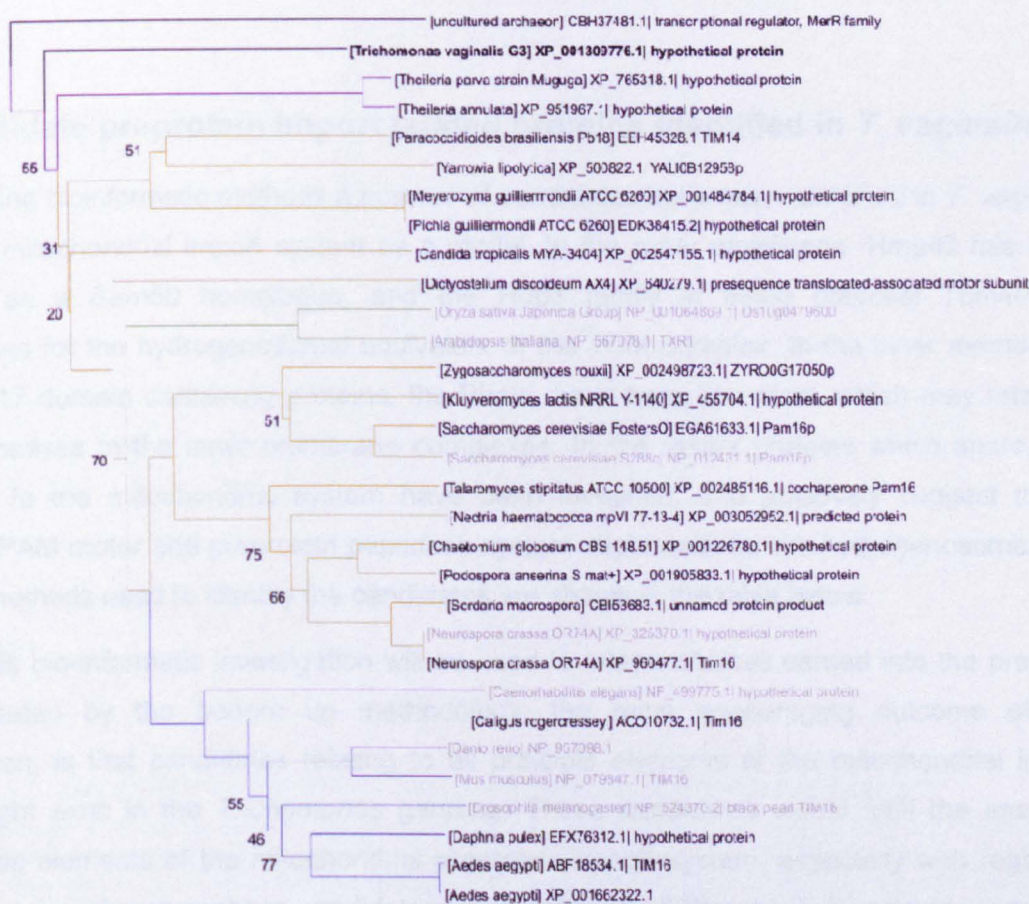


Figure 5.4.5 A phylogenetic analysis of the Tim23 peripheral protein Pam16 homologue in *T. vaginalis*.

The *T. vaginalis* Pam16 homologue (bold, black) was used in a BLAST using the BLOSUM45 matrix and a gap/extension of 15:2. Results which were $E < 0.2$ were collected (black), and aligned with selected Pam16 sequences (grey) based upon their annotation as Pam16 homologs. A neighbour joining tree was then constructed and bootstrapped 100 times, nodes where the bootstrap value is < 80 are shown.

The results of the phylogenetic analysis for Pam16 illustrate a more divergent protein as a cutoff of 0.2 had to be used to assemble a collection, however a similar distribution can be seen. The *T. vaginalis* protein clades deeply within the tree most closely to the protozoan parasite *Theileria*.

The inner membrane components examined in this section are far more conserved than other components examined in the bio-informatic investigation. The homology of these proteins makes clear that different components of the membrane translocase system have diverged at

different rates, and that different parts of the translocase system, might be reduced, or like the PAM motor, more or less complete.

5.5 Candidate preprotein import related proteins identified in *T. vaginalis*

Using bioinformatic methods a number of candidates have been identified in *T. vaginalis* using the mitochondrial import system as a model. In the outer membrane, Hmp43 has been indicated as a Sam50 homologue, and the Hup3 family in being potential Tom40 like translocases for the hydrogenosomal equivalent of the TOM complex. In the inner membrane, three Tim17 domain containing proteins, the Phats, have been identified, which may relate to the translocases in the inner membrane complexes. In the matrix proteins which share high homology to the mitochondrial system have been identified, and putatively suggest that a complete PAM motor and preprotein peptidase system might exist for the hydrogenosome. The different methods used to identify the candidates are shown in the table below.

This bio-informatic investigation will be used to inform choices carried into the practical work, dictated by the bottom up methodology, the most encouraging outcome of this investigation, is that candidates relating to all principle elements of the mitochondrial import model might exist in the *Trichomonas* genome. These candidates would fulfil the essential mechanistic elements of the mitochondrial preprotein import system, especially with regard to the translocases, however these candidates are significantly different in their primary sequence to other model organisms.

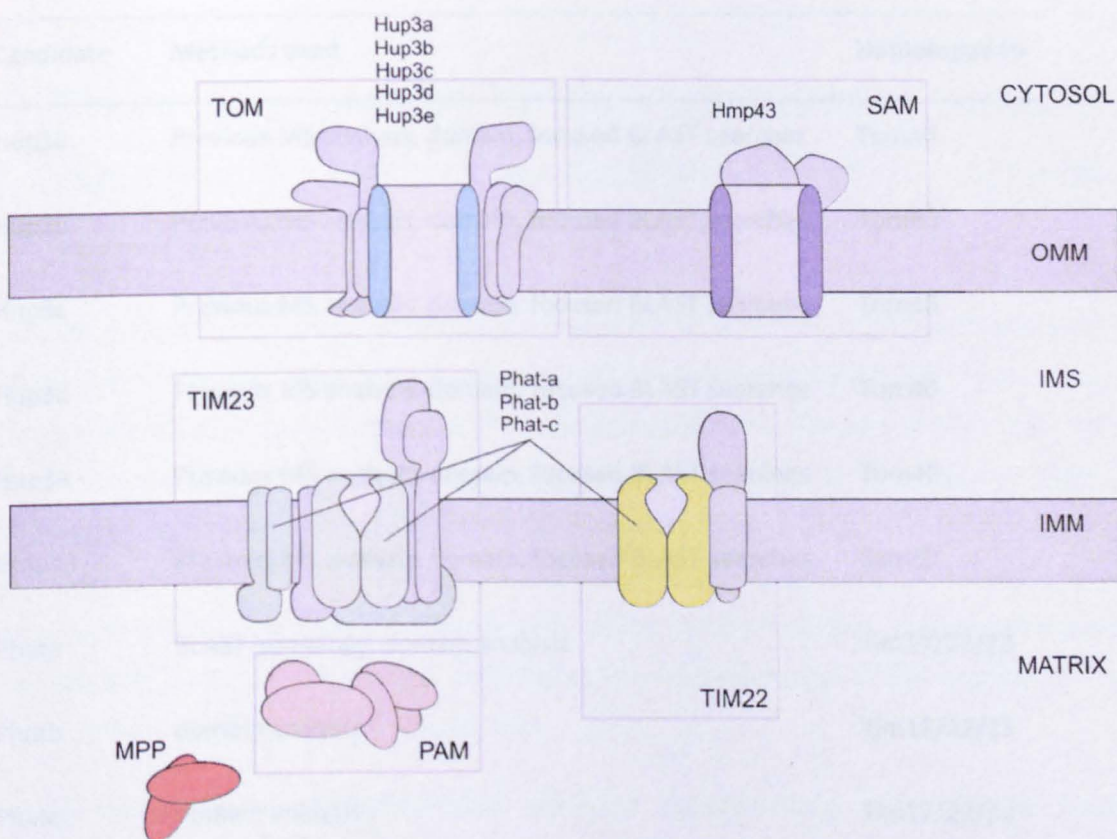


Figure 5.5.1 *T. vaginalis* homologue proteins super imposed of the mitochondrial import model

The bioinformatic approaches used in this chapter have identified a number of promising candidates for the hydrogensomal preprotein translocase system using the *S. cerevisiae* mitochondrial model. The Hup3 family are putative translocases in the TOM40 complex, whilst Hmp43 is a candidate for the SAM50 outer membrane complex. For the inner membrane the Phat family of proteins are candidates for the TIM22, 23 translocases. Finally convincing homologues for the PAM motor, mthSp70 and β -MPP have been found for the matrix compartment.

| Candidate | Methods used | Homologue to |
|--------------|--|--------------|
| Hup3a | Previous MS analysis, domain, focused BLAST searches | Tom40 |
| Hup3b | Previous MS analysis, domain, focused BLAST searches | Tom40 |
| Hup3c | Previous MS analysis, domain, focused BLAST searches | Tom40 |
| Hup3d | Previous MS analysis, domain, focused BLAST searches | Tom40 |
| Hup3e | Previous MS analysis, domain, focused BLAST searches | Tom40 |
| Hmp43 | Previous MS analysis, domain, focused BLAST searches | Sam50 |
| Phata | BLAST homology, domain analysis | Tim17/22/23 |
| Phatb | domain analysis | Tim17/22/23 |
| Phatc | domain analysis | Tim17/22/23 |
| β -MPP | BLAST homology, previously characterized | β -MPP |
| mtHsp70 | BLAST homology previously characterized | mtHsp70 |
| Pam16 | BLAST homology | Pam16 |
| Pam18 | BLAST homology | Pam18 |
| Mge1 | BLAST homology | Mge1 |
| Tim44 | BLAST homology | Tim44 |

Figure 5.5.2 Table illustrating the candidate proteins identified in the bioinformatic analysis, and the methods used in their detection.

The bottom up methodology is based on the detection of candidate proteins from bioinformatic investigation. To complete a putative system of translocases a variety different bioinformatic tools were used to support *T. vaginalis* proteins as candidates of this system. The table above illustrates the findings of this portion of the investigation, and the means through which they were identified.

6 Microscopic analysis of the ultrastructural features of *T. vaginalis* and localisation of recombinant hydrogenosomal proteins

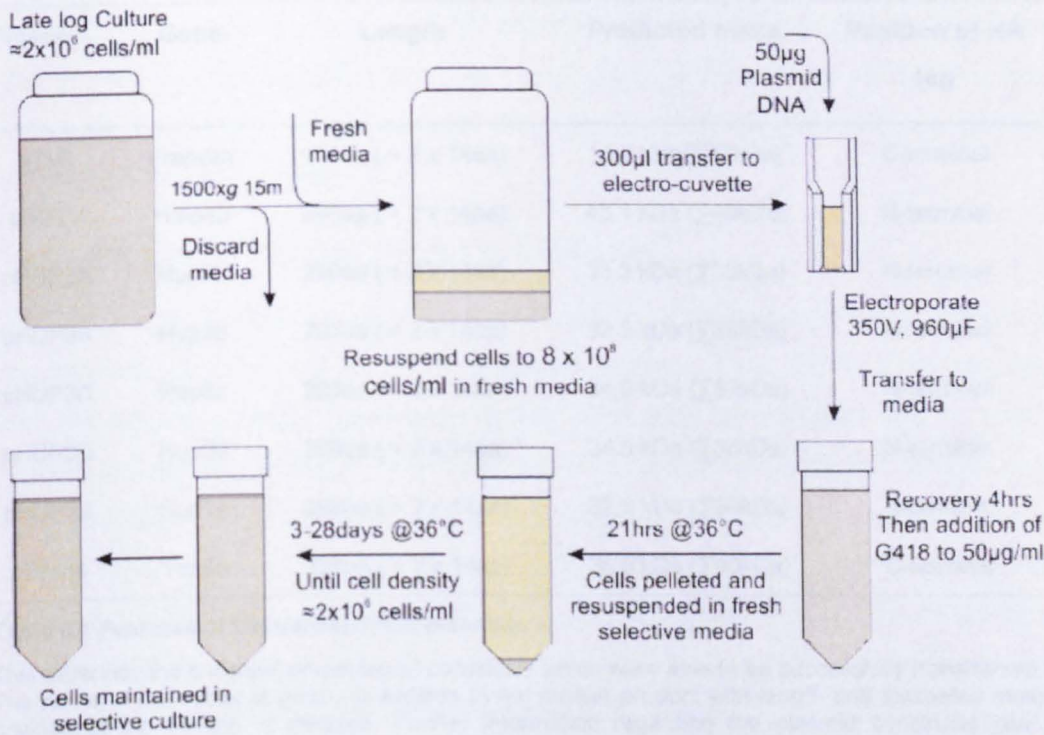
Chapter 5 identified a series of candidate proteins related to the mitochondrial import system. These candidates will move from *in silico* analysis to practical investigation by the production of *T. vaginalis* transformants expressing these candidates (6.1). These transformants will then be used in microscopic, and molecular biology investigations (Ch. 7). This section deals with the microscopy analyses and will examine both untransformed (6.3) and transformant (6.4) *T. vaginalis*, mention is also given to the techniques developed to study *T. vaginalis* microscopically (6.2). Whilst the investigation of the hydrogenosome, and candidate proteins is the foremost objective, microscopic techniques are also employed in this chapter to investigate other ultrastructural features of *T. vaginalis*.

6.1 Generation of transformant strains expressing HA tagged candidate hydrogenosomal proteins

The results of bioinformatic investigation were carried into laboratory research by the preparation of plasmids expressing HA-tagged protein. Special attention was taken to produce plasmids for the Hup3 family of proteins to represent the candidates for a putative TOM complex, Hmp43 for a putative SAM complex, and Phats a-c for inner membrane translocases TIM22, TIM23. In addition plasmids were prepared for the matrix protein homologues Tim44, and Pam16, as well as a control plasmid expressing the protein frataxin which was previously used to highlight mitochondrially derived organelles (Dolezal et al. 2007). In all cases the actual cloning required to generate the plasmids was carried out in the laboratory by Karen Lawler under direction from myself and Dr Dyall, and designed from the modified Topo I vector pTagVag, more information on p241. The sequence of each plasmid was verified prior to transformation of *T. vaginalis*.

Prepared plasmid was then used to transform the C1 strain of *T. vaginalis* (Methods p69, Figure 6.1.1), and 8 transformants were able to be successfully generated (Table 4.1). These transformants include the outer membrane candidates (Hmp43, Hup3a-e), and an inner membrane candidate (Tim44), as well as the frataxin control vector. Unfortunately no stable

lines were transformed expressing the Phat candidates, or Pam16, though purified plasmid was produced.



6.1.1 Generation of transformant *T. vaginalis*

The *T. vaginalis* strain C1 was used to generate candidate expressing transformants. Cells from dense cultures were pelleted and resuspended in fresh media to a density of 8×10^8 cells/ml. plasmid DNA was introduced and allowed to incubate on ice, before electroporation. Cultures were given 4hrs for recovery before the addition of G418, after 21hrs cells were resuspended in fresh media. Suspensions were allowed to grow until a density of 2×10^6 cells/ml, where upon they were maintained in selective culture. Full details of the method is given on p69

Table 4.1 highlights some of the essential features of the plasmid constructs. All candidate proteins were expressed in frame with a double HA-tag, which has been used for detection. The position of the HA-tag is not the same for each preprotein, this decision has been made to accommodate the natural processing of each preprotein, and to avoid tag cleavage in the mature protein. Frataxin has been shown to undergo signal sequence cleavage on maturation in both the mitochondrion and hydrogenosome (Gordon et al. 1999; Bradley 1997), and for the *T. vaginalis* homologues to frataxin and Tim44 the HA-tag has been positioned in frame with the C-terminal. The β -barrel proteins of the outer mitochondrial membrane are not known to be matured by signal sequence cleavage, but are known to have a highly conserved C-terminal signal (Kutik et al. 2008). It was then decided that in these constructs that the HA-tag be at the N-terminus, and fused in frame with the vector's start codon. Table 6.1 illustrates the

name of the vector, the protein product, its length and estimated mass, in addition to the position of the HA-tag.

| Plasmid | Gene | Length | Predicted mass | Position of HA tag |
|---------|----------|---------------------|----------------------------|--------------------|
| pTV2 | Frataxin | 121aa (+ 2 x 14aa) | 13.8 kDa(Σ 17kDa) | C-terminal |
| pHBTv | Hmp43 | 398aa (+ 2 x 14aa) | 43.1 kDa (Σ 46kDa) | N-terminal |
| pHUP3A | Hup3a | 290aa (+ 2 x 14aa) | 33.3 kDa (Σ 36kDa) | N-terminal |
| pHUP3B | Hup3b | 297aa (+ 2 x 14aa) | 32.5 kDa (Σ 36kDa) | N-terminal |
| pHUP3C | Hup3c | 305aa (+ 2 x 14aa) | 34.0 kDa (Σ 37kDa) | N-terminal |
| pHUP3D | Hup3d | 308aa (+ 2 x 14aa) | 34.8 kDa (Σ 38kDa) | N-terminal |
| pHUP3E | Hup3e | 296aa (+ 2 x 14aa) | 32.8 kDa (Σ 36kDa) | N-terminal |
| pTIM44 | Tim44 | 326aa (+ 2 x 14aa) | 36.5 kDa (Σ 40kDa) | C-terminal |

Table 6.1 Features of the transformant plasmids

This table lists the essential properties of constructs which were able to be successfully transformed into *T. vaginalis*. The name of the vector is given, in addition to the protein product with length and estimated mass. In addition the position of the HA-tag is denoted. Further information regarding the plasmid constructs can be seen in the supplementary data.

6.2 Preparation of fixed cells for microscopy

Prior to microscopic examination of transformant cells, approaches for fixing the cells were optimised. Scarce information was given in literature as to a reliable preparation technique, and two different approaches were tested before results were suitable for microscopic characterisation.

The first approach utilised methanol/acetone fixation, in combination with silane coated slides. However epifluorescent observation of the fixed cells indicated that they had taken a shrivelled appearance, in addition to very few cells being adhered to the substrate. This approach was superseded by a gentler method (Methods 2.12.2) whereby living cells resuspended in PBS were allowed to adhere to a silane covered slide. This approach had the benefit of binding the cells more effectively than the methanol/acetone approach. Bound cells were then permeabilised and fixed *in situ* on the slide. Whilst cell detachment was still an issue, especially through the length of an immunodetection study, more cells survived, and their appearance appeared was more natural under epifluorescence.

6.3 Imaging of non-transformed *T. vaginalis* C1

With an effective cell preparation technique established, cells from the C1 parent strain were examined to provide an ultrastructural baseline with which to compare the transformant strains. This opportunity was used to check the morphology of the cells to validate the fixation process, and to examine cellular compartments within cells, which can be used as a point of comparison for the transformant strains.

6.3.1 Confocal bright field images of fixed cells

Late log phase *T. vaginalis* C1 cells were fixed and mounted according to the protocol used in the methods (2.12.2) and were examined on a Leica C100 confocal microscope to examine their overall morphology (see Figure 6.3.1). Cell size corresponds well to prior observations (Benchimol 2009). Delicate structural features including axostyle, flagella and undulating membrane can be clearly seen. At high magnification and high resolution, the inner cell volume can be seen to be packed with punctuate organellar structures which crowd to the cell periphery. These images show that the preparation technique is able to conserve these delicate structural elements, and present cells which near normal phenotypic appearance. Example images from such cell preparation are shown for the Hup3C transformant in Figure 6.3.1 but similar images were obtained for untransformed *T. vaginalis*.

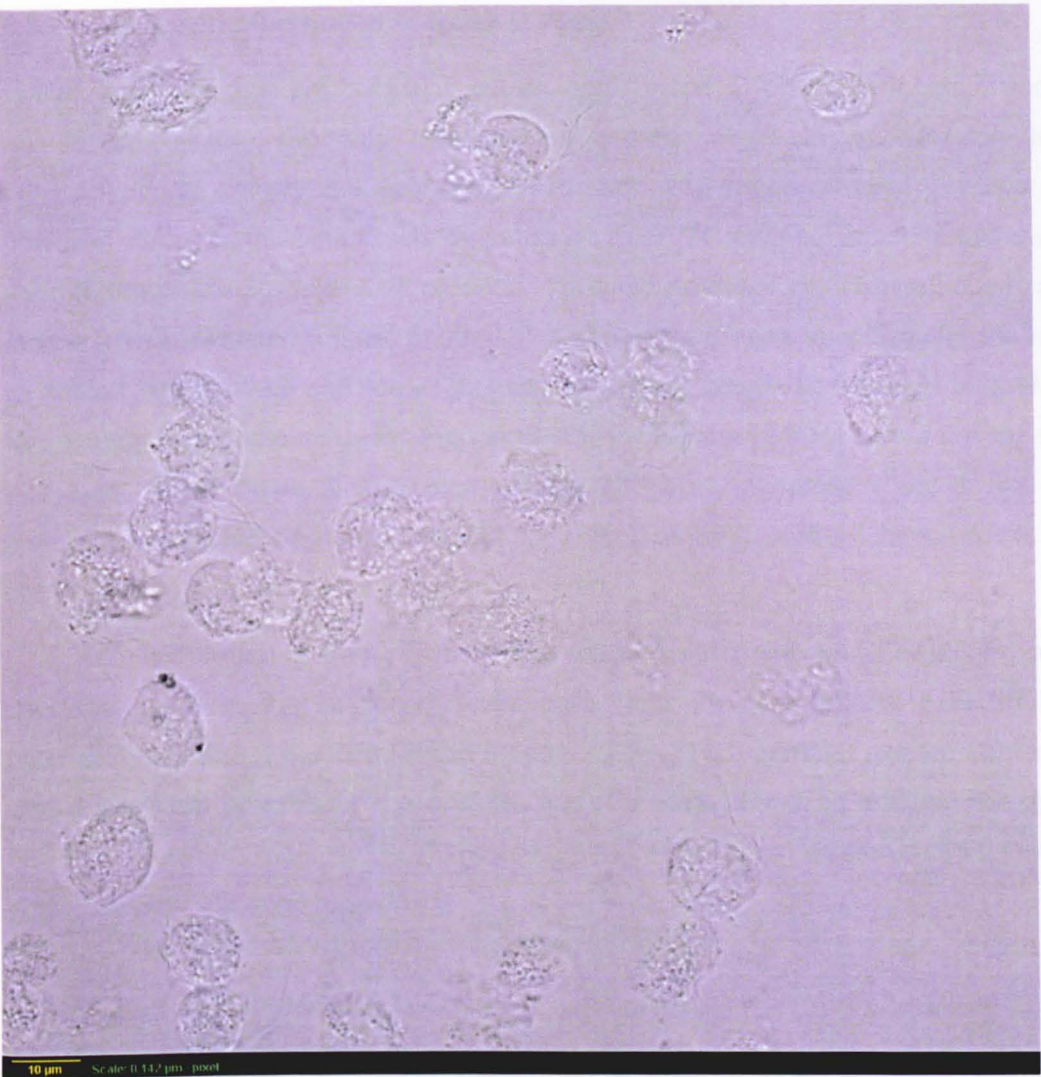


Figure 6.3.1 Fixed cells visualized by confocal transmission

Confocal imaging was performed on fixed Hup3c transformant cells. The Image above shows a confocal slice through fixed cells, illumination of the sample was achieved using a solid state 488nm laser, and the transmitted light was collected at the same frequency, by this approach a transmission type image was collected. The monochromatic light used for the imaging allows for much higher resolution images, the values of which are shown in the figure legend. The fixed cells show an abundance of structures which diffract the illuminating light, some of these features are highlighted in the following images.

6.3.2 Visualizing the nuclei of fixed *T. vaginalis* cells

In this study two different nuclear markers were tested, DAPI and Propidium Iodide (PI). Both agents are simple aromatic intercalation agents, which insert between base pairs in polynucleotides. Initially, the well used DAPI stain was tested on fixed C1 cells (Figure 6.3.2); however some signal “bleed” was detected on the FITC channel which would subsequently be used to image the transformant proteins. To avoid potential signal overlap a Propidium Iodide marker was substituted (Figure 6.3.3) which showed a greater specificity for the nucleus as well as limited signal bleed into the FITC channel. These observations were extended to confocal microscopy (Appendices: DAPI; Figure 10.4.1, PI; Figure 10.4.2), where the same effects were observed. Whilst confocal microscopy could effectively eliminate most of the cross channel bleed from the DAPI signal by stringent separation of the collection channels, the non-specificity of the DAPI localization is still prominently seen.

Consideration of these factors was used to determine that Propidium Iodide would be used for marking the nuclei of fixed cells from its compatibility with the epifluorescent microscope, where it exhibits limited bleed into the FITC channel, and for confocal microscopy where it is both specific for the nucleus and low background. In addition the use of PI would facilitate in using a wider range of visible light markers in combination as PI's emission region is in the red/near infra-red region.

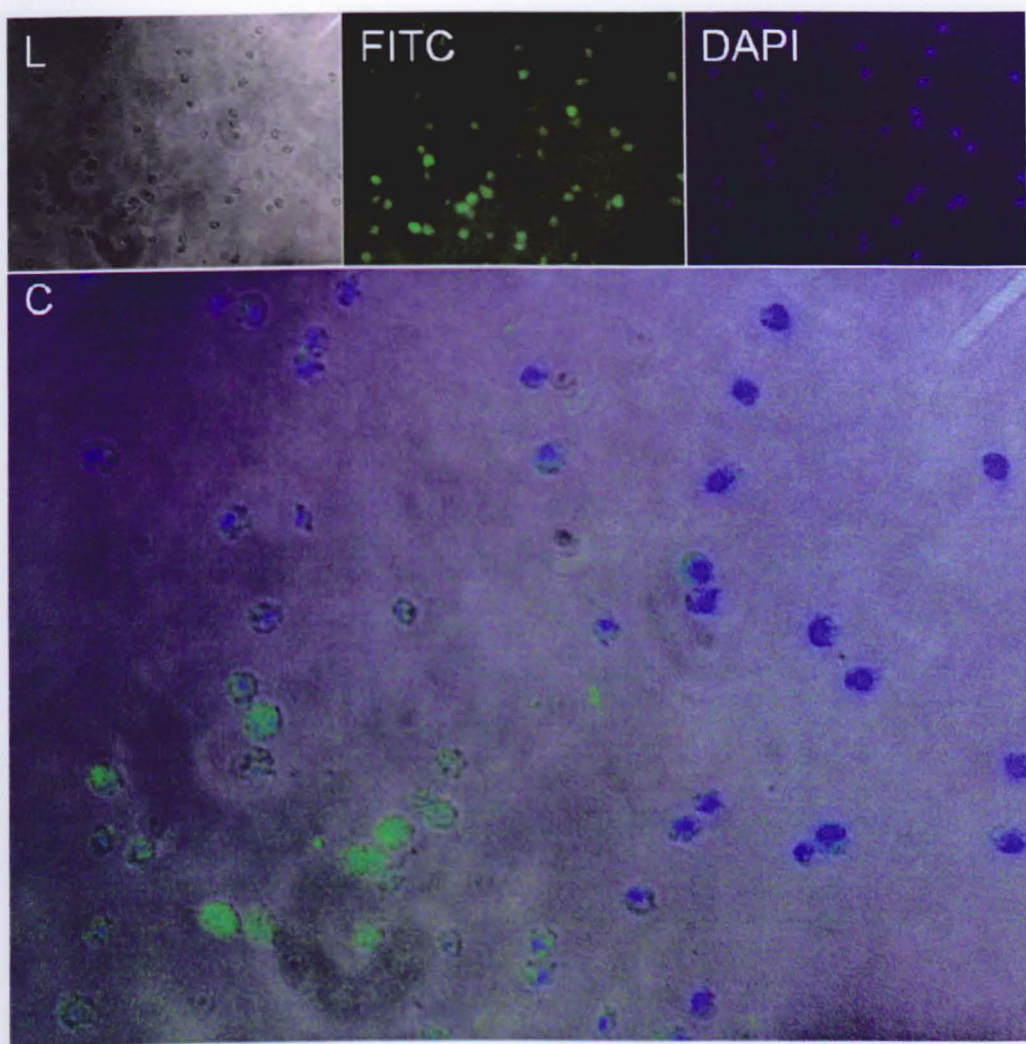


Figure 6.3.2 Nuclear localization in fixed *T. vaginalis* using DAPI counterstain

The nuclear stain DAPI was investigated for use in the fixed *T. vaginalis* cells. Fixed C1 were treated with DAPI and imaged on FITC and DAPI channels on an epifluorescence microscope in addition to white light image (L). The nuclear stain does focus on single large cellular bodies (seen in the DAPI channel) but does show a degree of bleed into the FITC channel. Additionally there is some degree of background in the image.

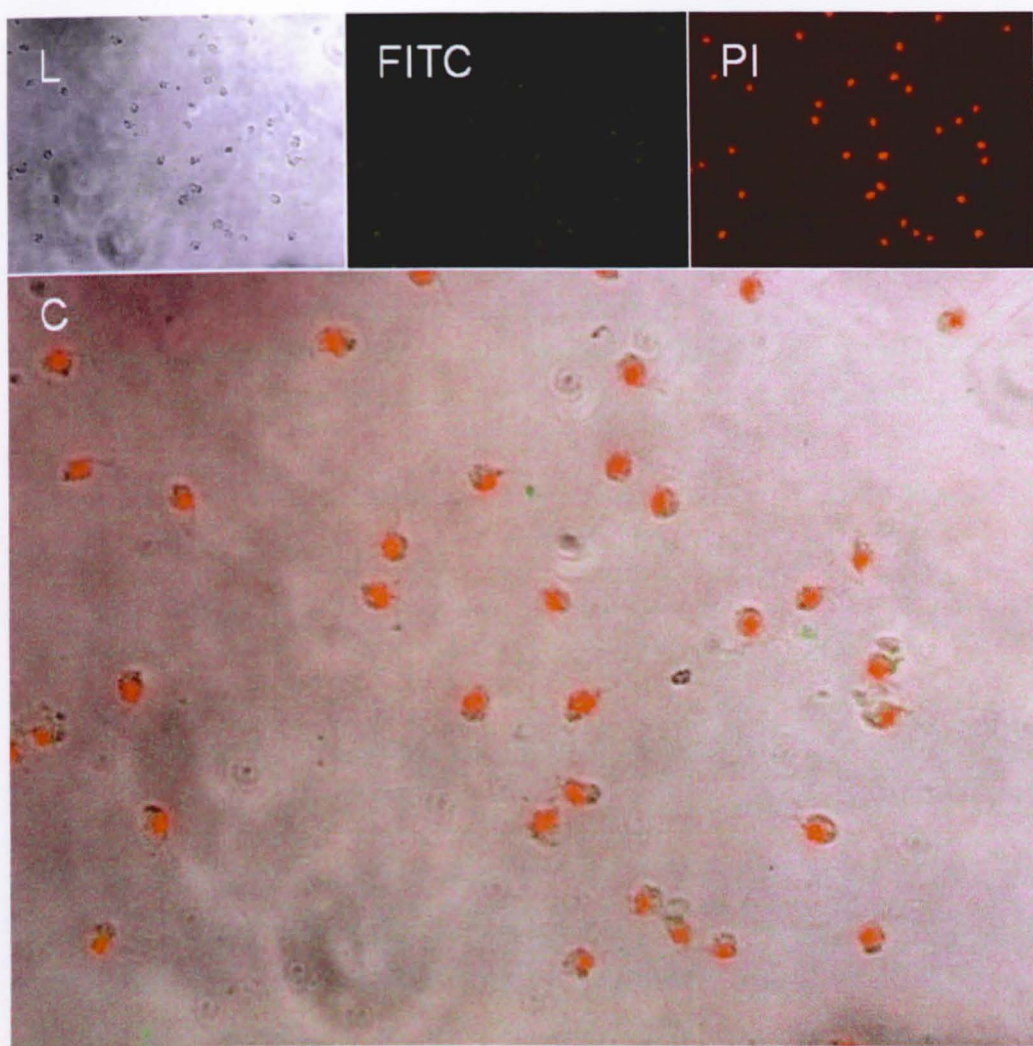


Figure 6.3.3 Propidium iodide better identifies the nucleus of *T. vaginalis*

The alternative nuclear stain Propidium Iodide was investigated on a epifluorescent microscope. The bright field image (L), and FITC and Rhodamine (PI) channels are shown. The longer emission wavelength of PI cannot be collected using the DAPI filter, and its emission is viewed here using the Rhodamine dye filter. The PI stained cells show tightly localized foci of intense staining which can be seen central to the cells in the combined image (C). What is more is that there is little bleed of the PI into the FITC channel, however there does remain some background apparent in the PI channel.

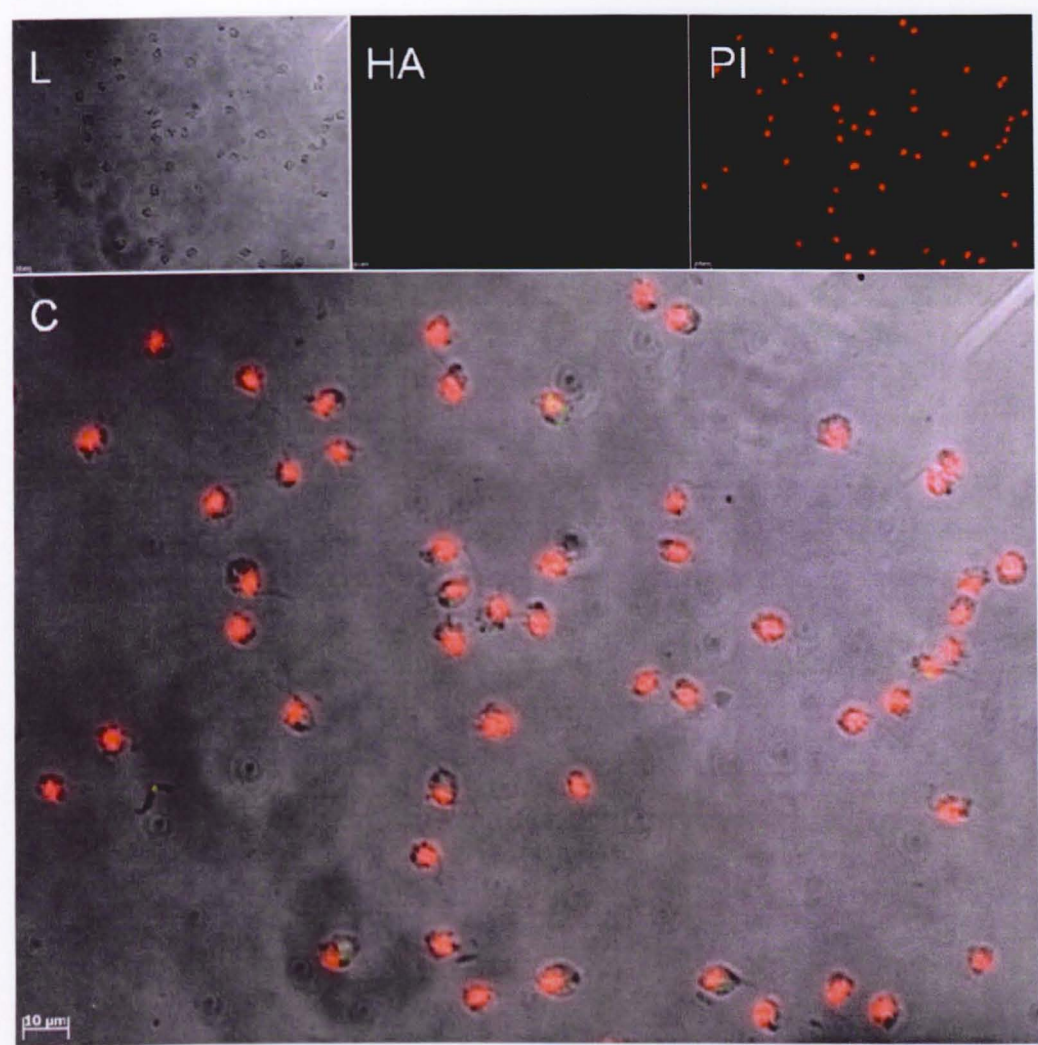


Figure 6.3.4 *T. vaginalis* cells incubated with propidium iodide and anti-HA antibody

Fixed C1 cells were incubated with a primary anti-HA antibody which was then detected using an Alexa488-fluor coupled secondary, which was imaged on the epifluorescent FITC channel (HA). The nuclear material of the cell was stained with Propidium Iodide and imaged on the Rhodamine channel (PI). A light image of the cells was taken (L) and the combined image of these channels can be seen in C. The PI stain demonstrates a tight specificity for the nuclei, with little extranuclear signal. The anti-HA body is not detected to a great degree in the fixed cells, indicating little latency/signal within cells lacking recombinant protein.

T. vaginalis cells were incubated with a primary anti-HA antibody which was then detected using an Alexa488-fluor coupled secondary, which was imaged on the epifluorescent FITC channel (HA). The nuclear material of the cell was stained with Propidium Iodide and imaged on the Rhodamine channel (PI). A light image of the cells was taken (L) and the combined image of these channels can be seen in C. The PI stain demonstrates a tight specificity for the nuclei, with little extranuclear signal. The anti-HA body is not detected to a great degree in the fixed cells, indicating little latency/signal within cells lacking recombinant protein.

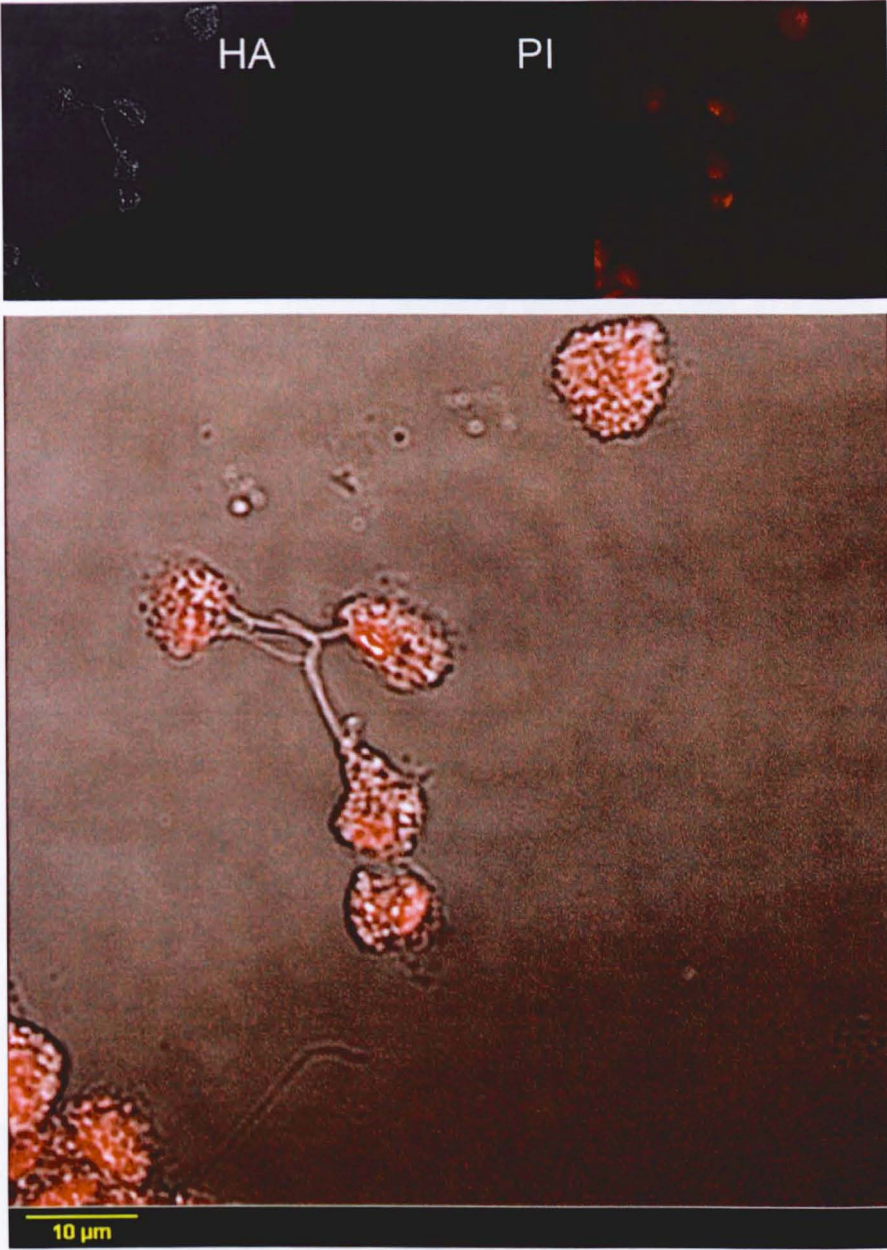


Figure 6.3.5 Distribution of the nuclear marker propidium iodide in fixed C1 cells

A composite image showing transmission (L), anti-HA (Detected by an Alexa488 coupled secondary antibody) and Propidium Iodide channel (PI). The transmission image was collected between 410-470nm, the Alexa488 emission at 498-547nm, and Propidium Iodide at 603-684nm. 405nm and 488nm lasers were used to illuminate the sample. The fixed C1 cells demonstrate very little antibody signal, seen before in the epifluorescent images, this would reinforce the evidence that this antibody generates little background signal in cells lacking transformant protein.

6.3.3 Visualizing the distribution of the ER compartment using an anti-BIP antibody

The distribution of the Endoplasmic Reticulum (ER) was explored in *T. vaginalis* C1 using an anti-BIP antibody (See Table 2.5 p77). Fixed cells were permeabilised and incubated with both anti-BIP antibody and PI. Confocal images are shown in Figure 6.3.6, and Figure 6.3.7.

The distribution of the PI nuclear signal is particularly clear in Figure 6.3.6 (shown channel 2). Three nuclei are clearly seen, and display the same kind of heterogeneous staining seen before with DAPI (Figure 10.4.1) The Zeiss confocal microscope allowed for excellent channel discrimination between PI and Alexa488 Channels, and no signal bleed is apparent.

The distribution of the anti-BIP marker is striking, and shows a crowded cell body packed with thousands of punctate signal sources. The confocal slice shown in Figure 6.3.6 cuts across the upper plane of the nuclei in the three cells, there is clear delineation of the nuclear volume seen in anti-BIP channel, though not all signal is entirely excluded from the nuclear region, this could either represent infiltration of the nucleus by structures which have become highlighted by the anti-BIP antibody, or that slice volume encompasses a region where both structures exist.

The majority of the anti-BIP signal is seen surrounding the nucleus, a 'halo' features seems to be apparent, this gap measures approximately one micron in width, and extends across confocal planes. This halo is not absolute however, and BIP signal can be seen immediately adjacent to the nucleus (cell on top right, and a clean point contact in the cell center bottom). Because PI and the anti-BIP tracking agents only partially reveal the full extent of their intracellular compartments it is not possible to resolve whether the halo gap represents a true barrier between these organelles, or whether it is a region of the nucleus or ER which simply does not generate signal with these markers.

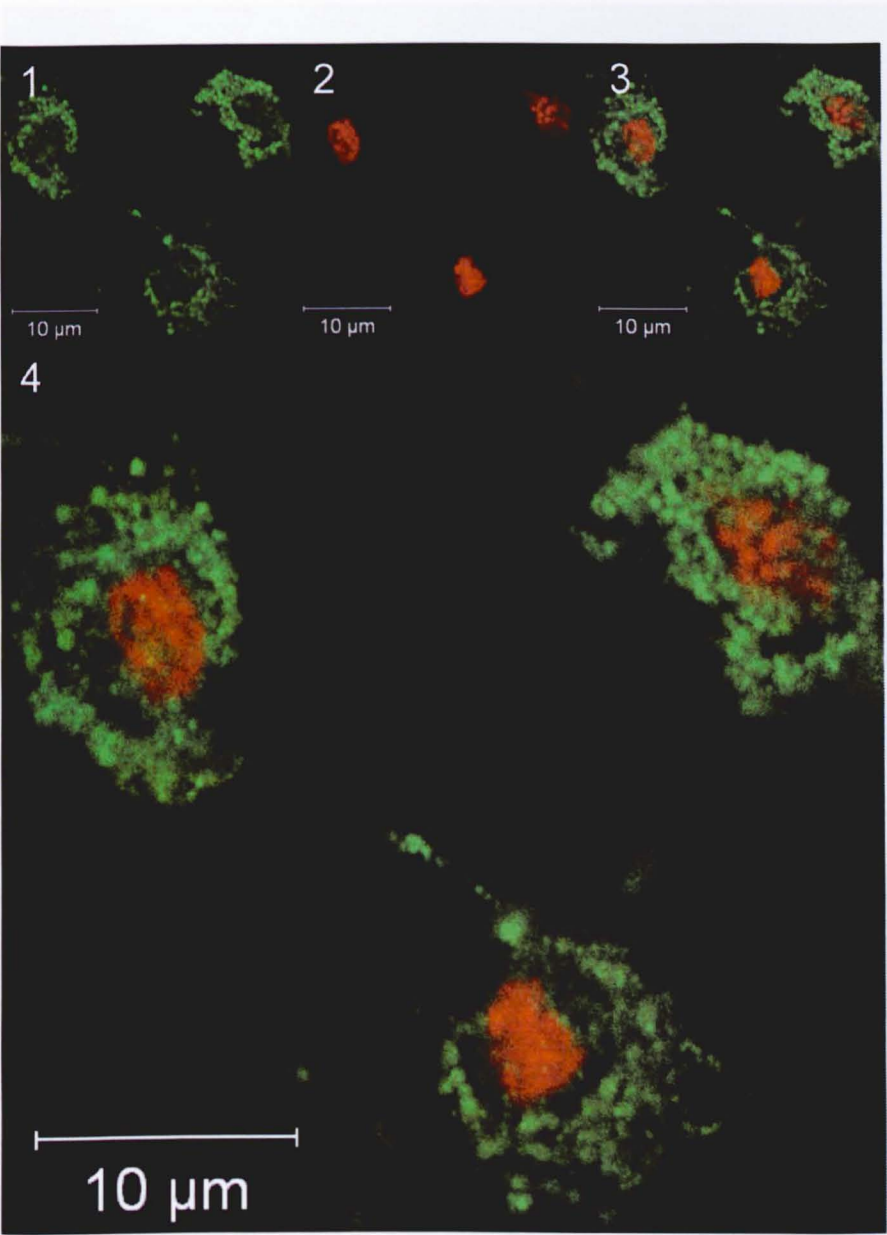


Figure 6.3.6 Distribution of anti-BIP in fixed *T. vaginalis* cells

The ER localizing anti-BIP antibody was used here in untransformed *T. vaginalis* C1 cells to show the distribution of ER (1), these cells were also incubated with the nuclear marker PI (2), the merged channels (3,4) show that the two localizations are discrete, and that there is a clear halo between the stained chromatin and the anti-BIP labeled ER. The anti-BIP antibody was detected using an Alexa488 secondary and was excited using a 488nm laser, light was collected for this channel between 493-555nm, the PI channel was excited using a 561nm laser and light collected at wavelengths of longer than 600nm. The area shown has a resolution of 0.066 $\mu\text{m}/\text{pixel}$.

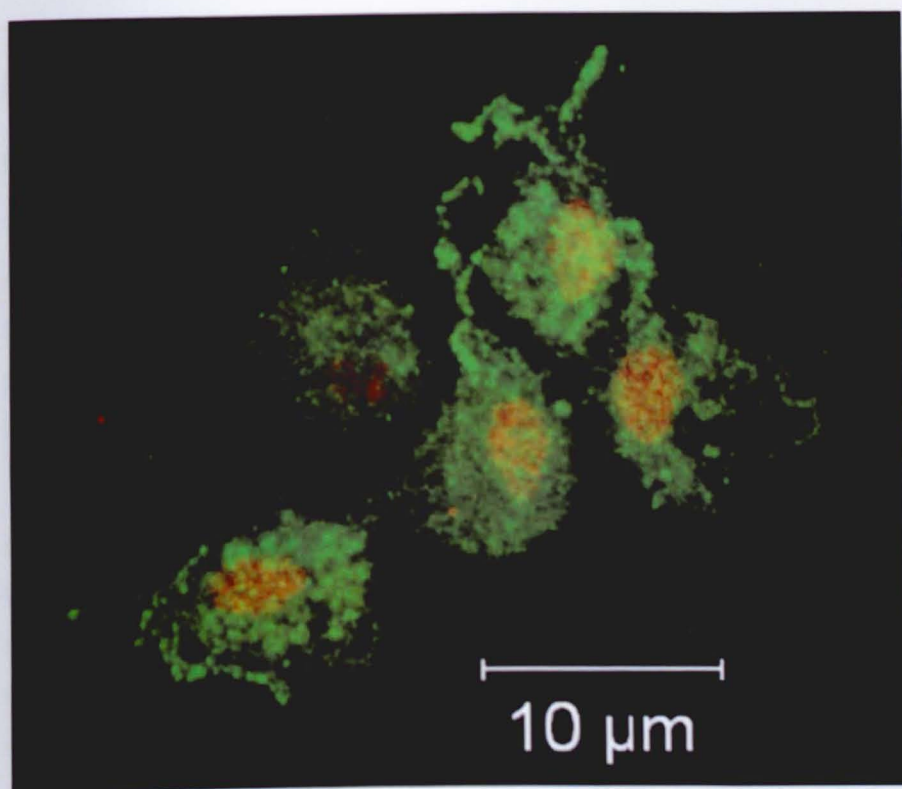


Figure 6.3.7 Distribution of BIP and PI trackers in reconstructed cells

A group of fixed C1 cells were imaged in a series of slices to a total volume $33.67 \mu\text{m} \times 33.67 \mu\text{m} \times 8.20 \mu\text{m}$ (z-axis) in twenty slices. The anti-BIP/Alexa488 complex was excited using a 488nm laser and light sampled between 501-530nm, this channel has been labeled in green. The PI marker was detected by exciting with a 561nm laser with light collected between 570nm-600nm and shown in red.

Anti-BIP signal is concentrated into myriad punctate sources. This type of localisation pattern is consistent with the distribution of BIP seen in other investigations (as in closely related *Giardia*: (Stefanic et al. 2006)). These punctate sources vary from near single voxel ($\approx 100\text{nm}$) to $0.5\mu\text{m}$ (500nm) structures, though average size is $\approx 0.3\mu\text{m}$. Given the dimensions of an antibody fluorophore complex are $>20\text{nm}$ these sources represent clusters of protein detected by the anti BIP primary, even at the smallest scales.

Whilst anti-BIP signal is detected in punctate sources there is an apparent higher order in which these sources appear to be arranged. Some evidence of this is seen in Figure 6.3.6, however it also shows some localisation to flagellar structures, this is more clearly seen in Figure 6.3.7. It is unclear why anti-BIP signal is detected in these structures which are assumedly unrelated to the ER. Most anti-BIP signal is related to the cell body, Figure 6.3.7 shows this particularly clearly- illustrating that the anti-BIP signal is arranged in a dense cloud surrounding the nucleus. Some signal can be seen arranged as ribbons, with punctate sources distributed along their lengths.

This would seem to indicate that whilst the BIP signal might be concentrated to discrete foci, that the BIP is part of a larger cell spanning reticulate structure consistent with ER.

6.3.4 Visualizing Golgi structures using Invitrogen molecular probes NBD C6-ceramide

The Invitrogen molecular probes NBD C6-ceramide marker, was used to identify the Golgi compartment, this probe was incubated with live cells to allow for dye accumulation, before fixation. In Figure 6.3.8 cells were incubated with the Golgi marker, and were subsequently processed with PI after permeabilization. The PI channel (1) shows the same consistent PI localisation as seen previously, with some diffuse signal in the cell body.

The Golgi marker signal is shown in channel 2 and shows a more diffuse distribution, but is none the less confined to within the cell body, and not in the intercellular space in the image. Whilst diffuse it does still show several distinct features. First is a nuclear void which tightly outlines the nucleus stained with PI (in contrast to the nuclear halo seen for BIP in Figure 6.3.6). Colocalization between the Golgi marker and PI is very limited. In addition to the nuclear void, several structures show increased signal intensity over background.

The Golgi marker is particularly intense around a single flattened perinuclear body and extends above and below this confocal plane, in constant position with respect to the nucleus. The size and position of this body would be consistent with the *T. vaginalis* Golgi body structures previously described (Benchimol et al. 2001). The Golgi marker is also seen intensified in a membrane surrounding the nucleus, and singly distributed micron sized bodies scattered in the cytoplasm, which are limited to only a few confocal slices. The identity of these structures is not apparent, however from the membrane inserting nature of this tracker we can conclude that these bodies have inserted into membranous structures. This would also explain why the signal is so poor within the nuclear region, and the diffuse labeling of the cytoplasm which is likely crowded with membranous structures as partly indicated in Figure 6.3.6 from the distribution of the ER marker. Figure 6.3.9, Figure 6.3.10 Figure 10.4.3 show 3-dimensional reconstructions of the Golgi body within fixed cells. The weak nonspecific signal of the marker in these reconstructions has revealed additional aspects of the cellular ultra-structure, including flagella, axostyle, and undulating membrane.

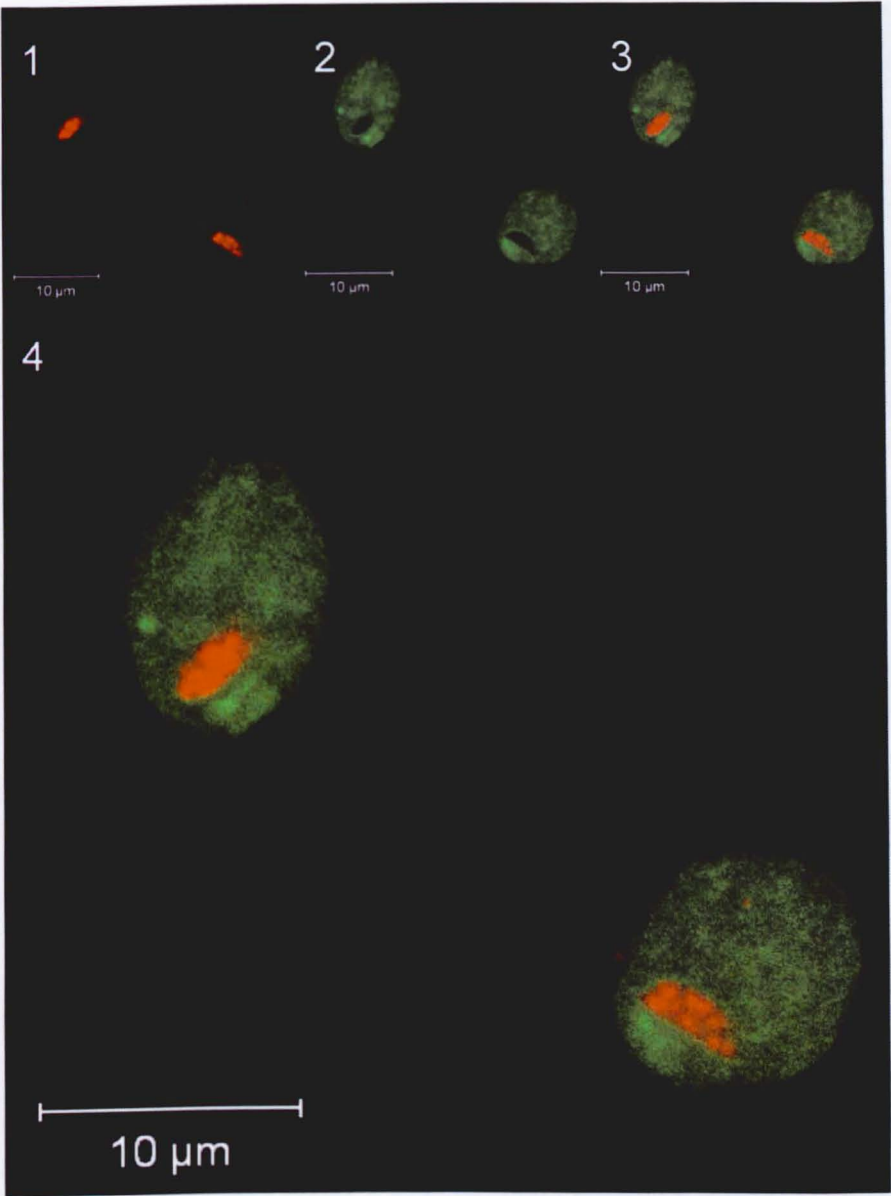


Figure 6.3.8 Distribution of PI and Golgi marker in fixed cells

Fixed untransformed *T. vaginalis* C1 cells treated with Golgi marker NMD-C6-ceramide (2), and nuclear marker PI (1). Combined images (3,4) reveal that the Golgi marker has a rather diffuse distribution within the fixed cells, but exhibits a concentration adjacent to a nuclear void. The image shown above is a 33.67 μm wide slice, with a resolution of 0.066 $\mu\text{m}/\text{pixel}$. The Golgi dye was excited at 458 nm, and emitted light collected between 494–540nm, the PI was excited using a laser at 561 nm, and wavelengths collected above 600nm.

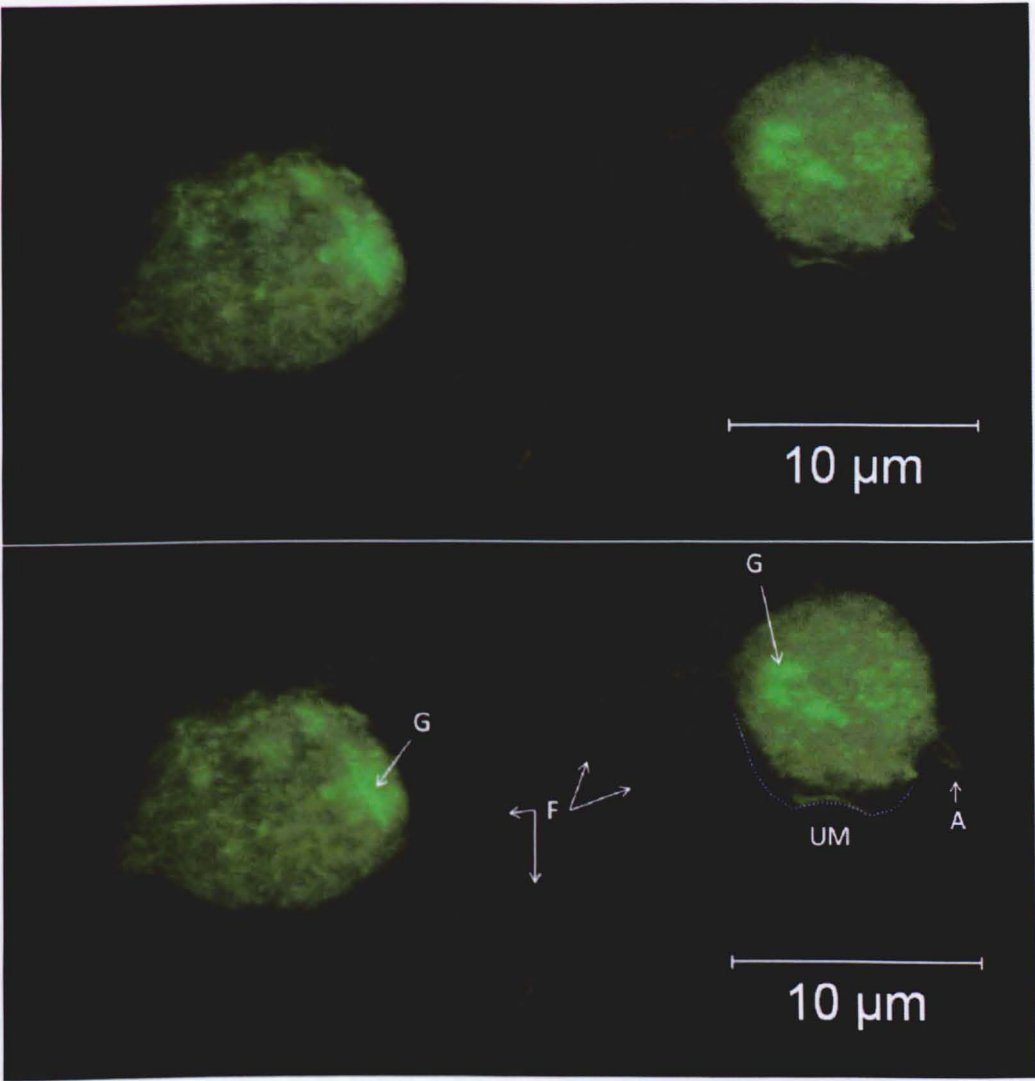


Figure 6.3.9 A 3D reconstruction of a Golgi apparatus

Fixed C1 cells stained with the Golgi marker were subject to confocal microscopy, a volume of 33.67 μm x 33.67 μm x 15.91 μm (z-axis) was sampled in 39 slices. These slices were then used to build a reconstruction shown above. The dye was excited using a 458nm laser and light was collected between 482-695nm. The lower image shows annotation of the reconstructed cells, indicating Flagella (F), Axostyle (A), and Golgi (G), the extents of the Undulating Membrane are shown as a dashed line (UM).

To investigate the distribution of the lysosomal compartment, the lysosomal molecular marker, LysoTracker (LTD 10) was included with the cells during fixation. Some issues were encountered with the use of LysoTracker in *Trichomonas*. Signal intensity was limited at the recommended concentrations of 100, especially the emission wavelength of the LysoTracker proved to be incompatible in combination with 21 nuclear marker. Figure 6.3.11 shows a two channel image (3) consisting of a bright field image (2) in combination with the LysoTracker channel (1). In addition a 3D reconstruction of just the LysoTracker channel is seen in Figure 6.3.12.

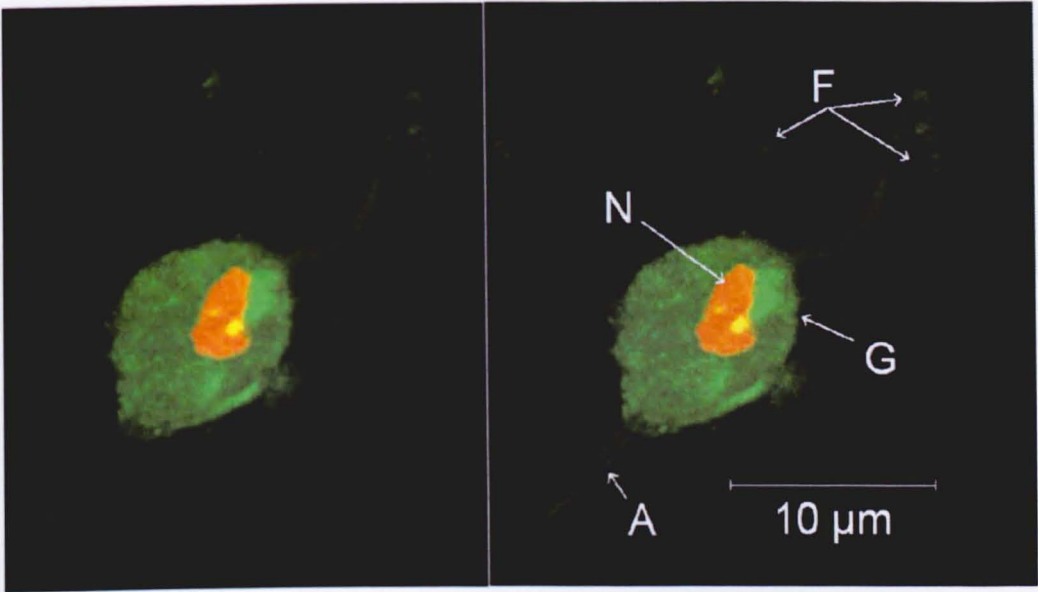


Figure 6.3.10 3D reconstruction of a Golgi apparatus and nucleus

A 3D reconstruction of a single fixed *T. vaginalis* C1 cell is shown above, the cell was imaged using the Golgi marker (Green) excited with a 458nm laser with emitted light collected 471-527nm, and the PI nuclear marker (Red), excited with a 561nm laser and emitted light collected at >600nm. The right hand image is shown annotated for Flagella (F), Axostyle (A), Golgi (G) and Nucleus (N).

The reconstruction of these structures corroborates that the fixation process is gentle enough to preserve these structures, and provide a natural looking cell for imaging. These reconstructions also enable the measurement of the dimensions of the Golgi body, the structure is clearly of a flattened structure, its plane contiguous with the nucleus, and the depth of this flattened plane is 1-2μm, and its length and width 4-5μm. These dimensions are consistent with previous work performed in *T. vaginalis* (Benchimol et al. 2001).

6.3.5 Visualizing acidic membranous compartments with Invitrogen LysoTracker.

To investigate the distribution of the lysosomal compartments the Invitrogen molecular probes marker LysoTracker (DND-99) was incubated with live cells before fixation. Some issues were encountered with the use of this tracker in *Trichomonas*, signal intensity was limited at the recommended concentrations of use, additionally the emission wavelengths of the tracker proved to be intractable in combination with PI nuclear marker. Figure 6.3.11 shows a two channel image (3) consisting of a bright field image (2) in combination with the LysoTracker channel (1). In addition a 3D reconstruction of just the LysoTracker channel is seen in Figure 6.3.12.

The tracking dye is diffusely distributed within the cells, however some punctuate ≈ 0.5 -1.0 micron sized structures can be discerned from background. In Z-stack imaging these structures do not persist beyond 2-3 slices indicating they are small topological isolated structures, their distribution does not appear to relate to the axis of cell (See Appendix 23 Intracellular distribution of LysoTracker bright bodies p261). In addition only between 3-6 of these bodies are seen per cell, however all cells show these structures, but with little pattern to number of localisation within the cell body.

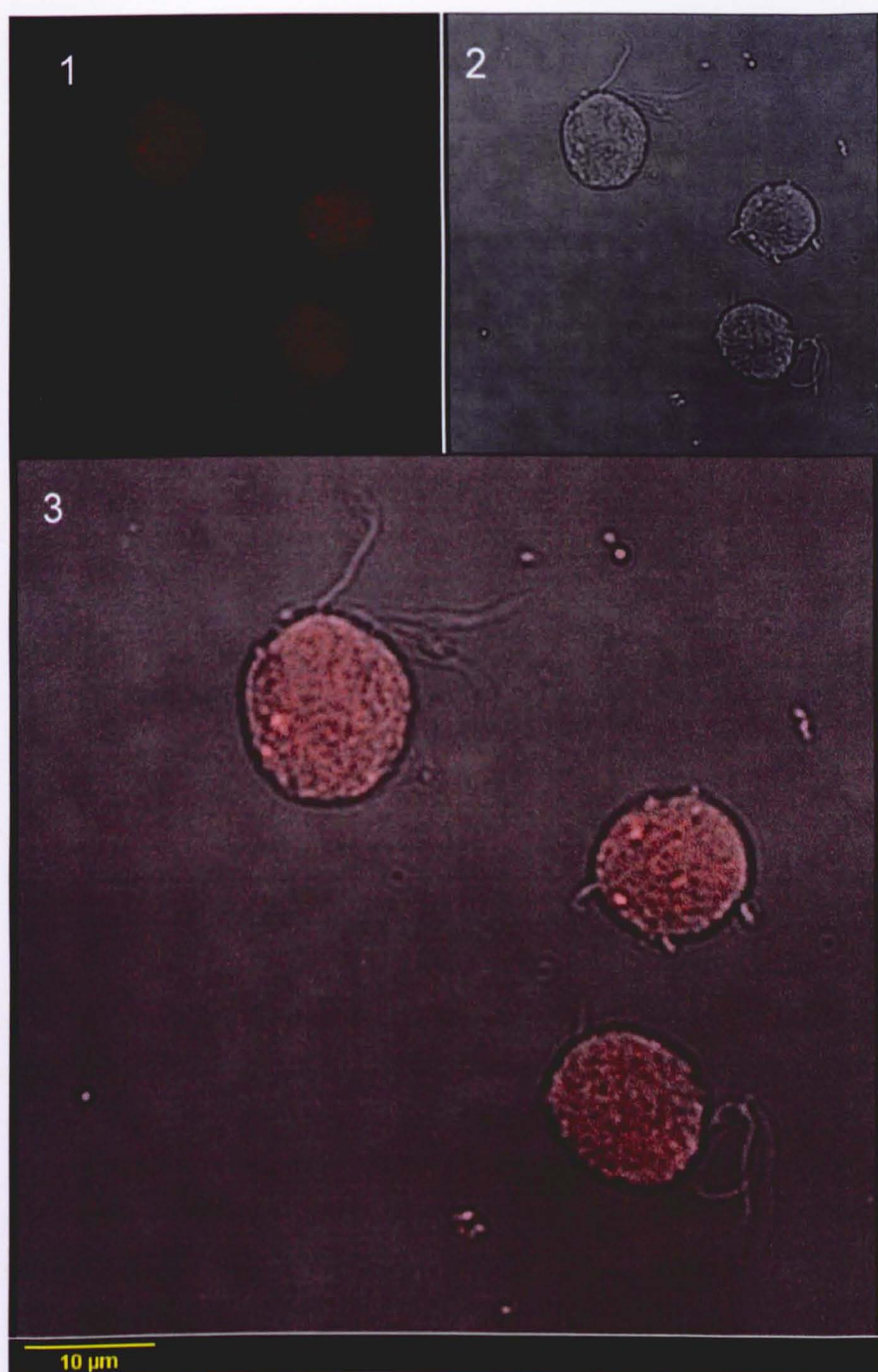


Figure 6.3.11 Distribution of LysoTracker dye marking acid vesicles

T. vaginalis C1, were pre-incubated in LysoTracker (DND-99), before fixation with paraformaldehyde minus Triton X-100. Confocal transmission image (1) and LysoTracker fluorescence (2) reveals that the dye signal is diffusely concentrated within atypical vesicles (3). The width of the image is 66.2 μm imaged at a resolution of 0.13 $\mu\text{m}/\text{pixel}$, the scale is shown at the bottom of the figure. The sample was illuminated using a 561nm laser which was used to compose the transmission image, the LysoTracker dye was sampled between 587-610nm.

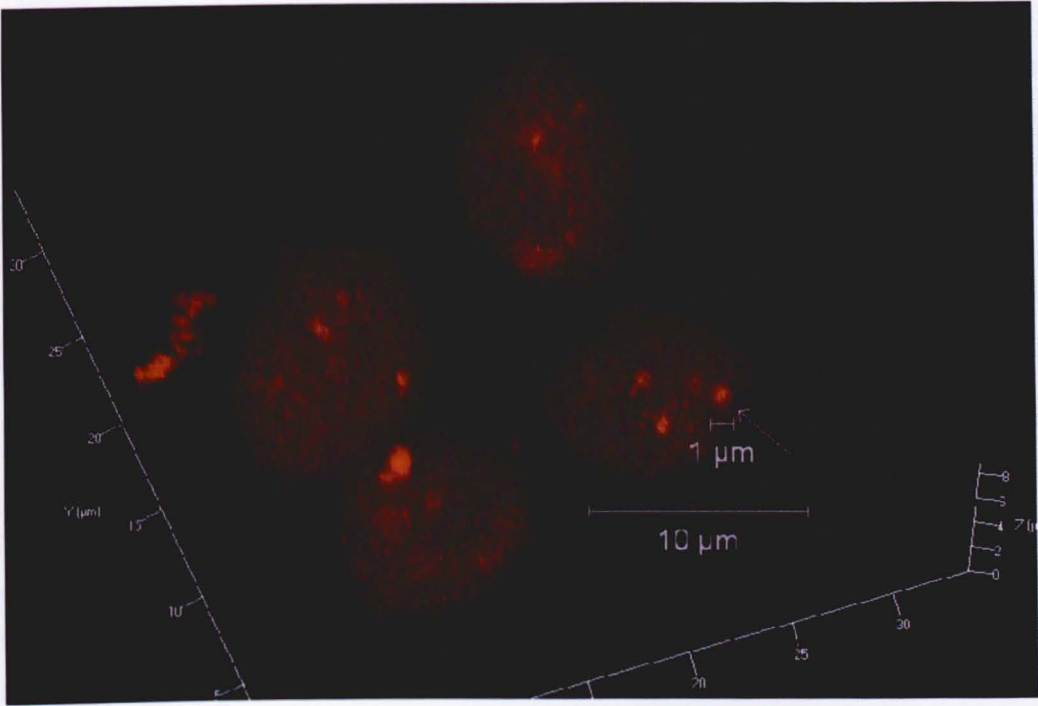


Figure 6.3.12 A reconstruction of fixed cells showing distribution of LysoTracker marker

Confocal sections were taken of a group of fixed C1, 21 slices of $33.67\ \mu\text{m} \times 33.67\ \mu\text{m}$ spanning a depth of $8.86\ \mu\text{m}$ were taken. The LysoTracker dye was excited using a 561nm laser, and fluorescence sampled at $566\text{-}690\text{nm}$. The images were then reconstructed and shown above. LysoTracker signal is concentrated in $\approx 1\ \mu\text{m}$ sized bodies, spanning multiple slices.

LysoTracker has similar membrane inserting properties to the Golgi ceramide marker seen in Figure 6.3.8, but its intensity strongly correlates to compartment pH. The brighter foci seen in Figure 6.3.11 may not reflect regions of increased dye concentration, but rather acidic compartments, where the pH exaggerates signal intensity over dye concentration. The signal intensity for the LysoTracker dye is insufficient to deduce clearly the compartments being identified. The size, and inferred acidity of these compartments would be consistent with lysosomes or peroxisomes, however these might be expected to be much more numerous. The larger bodies, and their limited number might also infer another type of discrete acidic compartment, perhaps similar to rhoptry in apicomplexans, an acidic organelle with an enzyme secretory role, which might reconcile the known conditional protease excretory activities of *Trichomonas*.

The collection of images shows the structure of the *T. vaginalis* cell, and the distribution of nuclear, ER, Golgi and lysosomal trackers. This information will be used to compare against the localization patterns for the HA-tagged recombinant proteins. In addition these works demonstrate that compartment investigation using immunodetection and confocal microscopy can provide insights into cell biology without recourse to electron microscopy.

6.4 Localisation of HA-tagged protein within hydrogenosomes

The localisation pattern of the plasmid expressed HA-tagged candidate proteins, were explored using a primary anti-HA antibody, with an Alexa488 coupled secondary antibody.

Some analysis of the localisation patterns were initially obtained under epifluorescent microscopy, which can be seen in Figure 10.4.4 p254, and Figure 10.4.5 p255. Whilst a mutual exclusion of HA-tagged protein and nuclear marker could be easily resolved, the microscope was not powerful enough to resolve the fine details of the localisation, this decision dictated a move to confocal imaging, which will be shown from this point onwards for detailing the localisation patterns of the transformant proteins.

6.4.1 Localisation of frataxin to the hydrogenosome

The soluble matrix protein frataxin, was used in this study as a control for the identification of the hydrogenosomal compartment. This approach was previously used in the work by (Dolezal et al. 2007) and provides an excellent marker for the hydrogenosome. frataxin also represents a unique probe for exploring the hydrogenosomal matrix, whereas the transformants which were able to be produced all have membrane associated roles.

Cells from the transformant line containing the HA-tagged frataxin were fixed and treated according to the methods described previously (Methods 2.12.2). A composite image was produced from the PI channel (1) and the fluorescently labelled antibody (2) which are also shown combined (3,4) in Figure 6.4.1

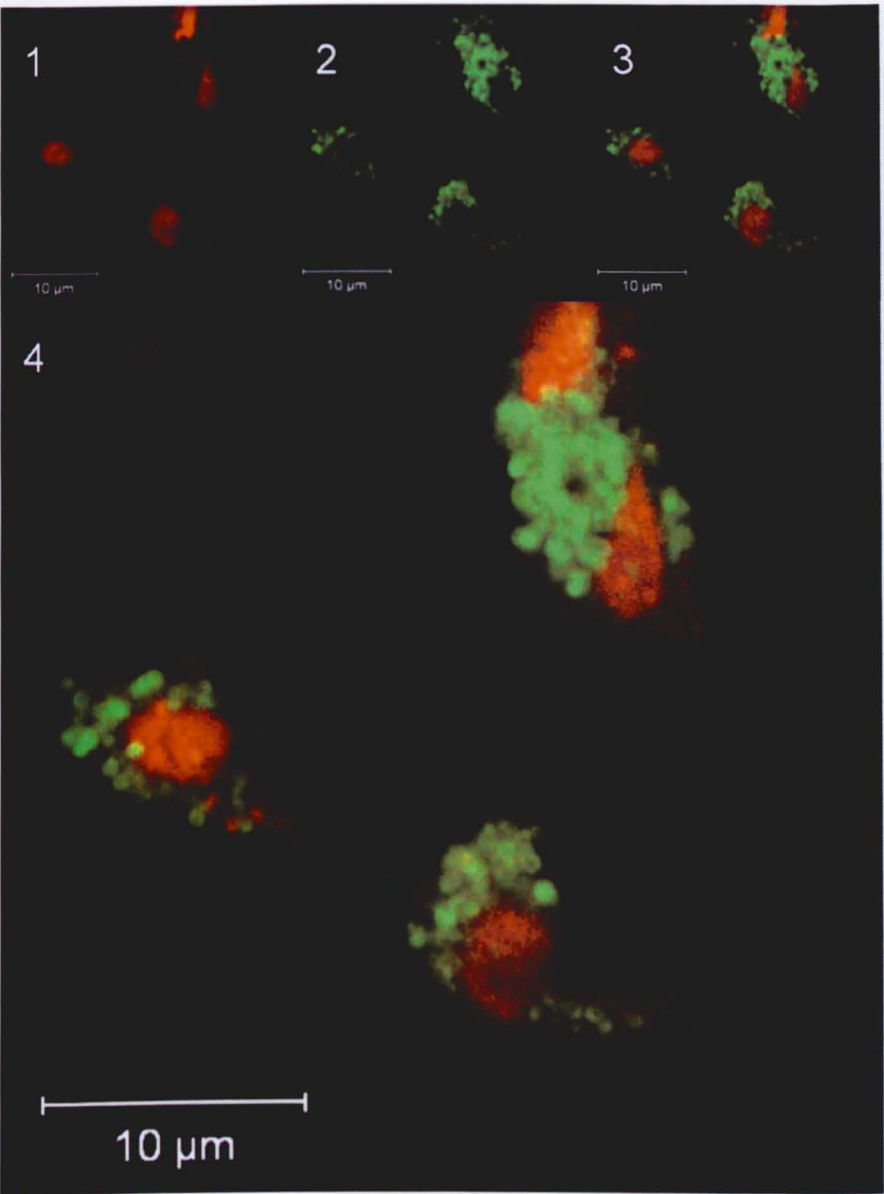


Figure 6.4.1 Intracellular distribution of frataxin

Frataxin-HA expressing transformant *T. vaginalis* C1, were fixed and stained with nuclear marker PI (1), and fluorescently coupled anti-HA/Alexa488 antibodies (2, green). The combined images (3,4) reveals frataxin localization to rounded bodies scattered within the cytosol consistent in size with the hydrogenosome. At high resolution (4) these are revealed to have a mix of rounded circles or ringlets. The Alexa488 conjugate was excited using a 488nm laser and emitted light collected between 493-555nm, the PI was excited with 561nm laser and light collected between 600-633nm. The width of the image is 33.6 μm with an imaging resolution of 0.066 μm .

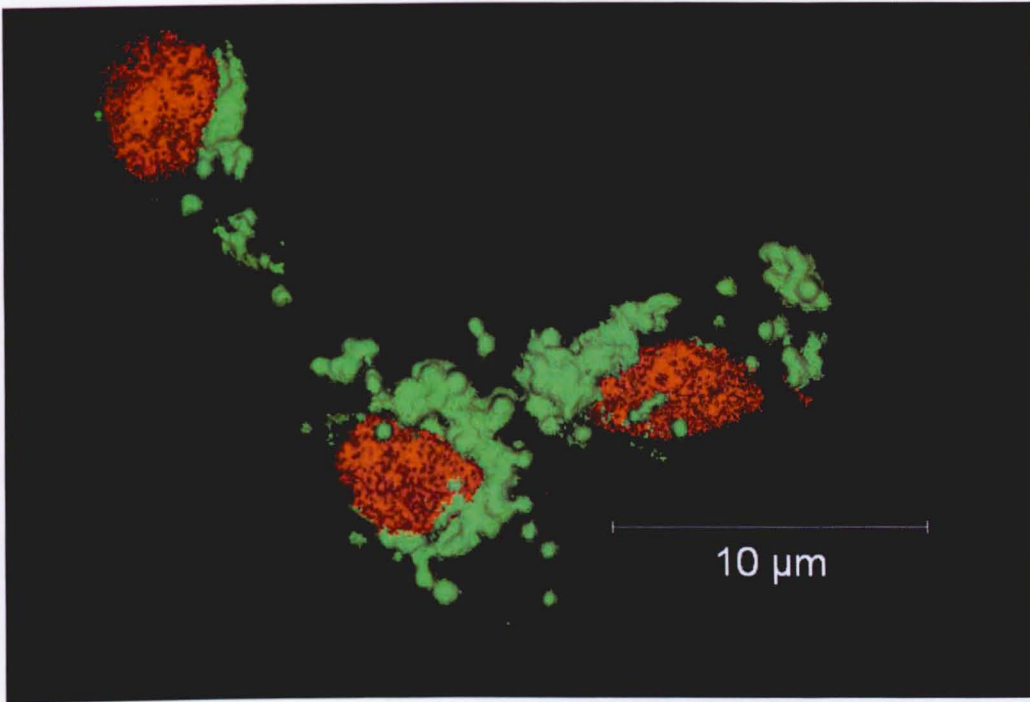


Figure 6.4.2 Reconstruction showing the structural arrangement of hydrogenosomes

Fixed C1 frataxin transformant cells were treated with PI (red) and Alexa488/Anti-HA (green), this image shows the result of a reconstruction performed on a series of 10 slices, in the ZEN-LE program. In this image the program has joined regions of equal intensity, creating an iso-surface. The resolution of this reconstruction is limited to the voxel resolution ($0.066\ \mu\text{m} \times 0.066\ \mu\text{m} \times 0.636\ \mu\text{m}$)

The distribution of HA-tagged protein is seen to distribute in a new distribution pattern compared to markers so far examined, signal is detected in clearly round organelles ≈ 1 micron in diameter. These bodies are consistent with the structures seen for frataxin in the work by (Dolezal et al. 2007) and the spherical bodies seen for *T. foetus* (Benchimol 2009), and which we can infer that the HA-tagged frataxin in Figure 6.4.1 is localised to the hydrogenosome. Using an edge detection method in reconstructed cells (Figure 6.4.2, Figure 10.4.6 p256), the spherical, and topologically isolated nature of the hydrogenosomes can be clearly seen.

The size distribution and population of hydrogenosomes are more thoroughly investigated later (Appendix 10.4.11 p262), but the characteristic, large, round, regular organelles seen with frataxin here are a clearly different to ER, Golgi or Lysosomal markers. There is some evidence of nuclear halo between the edge of the PI labelled nuclear material and the hydrogenosomes, though this is not as pronounced as the halo seen for the anti-BIP marker (Figure 6.3.6). The Golgi volume seen in Figure 6.3.8 should cause some displacement

of hydrogenosomes, though this is not clearly shown in this figure, however later figures do indicate a 'Golgi void'.

Frataxin is a soluble matrix protein and should occupy the inner volume of the hydrogenosome, some indication of this is seen in the image where there are hydrogenosomes which give a uniformly bright signal across the section, however there are subpopulations where the frataxin signal is concentrated at the organelle periphery, though it is unclear whether the signal is concentrated at the periphery of the matrix, inter-membrane space or hydrogenosomal membranes.

No de-convolution was performed on these images, and signal intensity reflects fluorescence at the confocal plane, meaning that the circlet features are not an artefact of image processing but represent real fluorescence. An alternate explanation could be that a subpopulation of hydrogenosomes have complex internal topologies. The existence of hydrogenosomal peripheral vesicles have been characterised in electron microscopy sections of the hydrogenosome, in *T. foetus* (Benchimol 2000; Benchimol 2008). Invaginations of the inner membrane would displace the matrix compartment into the ring shapes seen in the confocal slice, However if this was so then one would also expect to see crescent shapes, which are not clearly indicated in this image. Whilst peripheral vesicles would have a pronounced effect on matrix volume, the distribution pattern seen could arise by other effects.

Whilst the intra-organellar distribution of frataxin is hard to explain, the structures revealed are consistent with the hydrogenosome, also signal appears to be limited to this compartment, with very little signal outside of these organelles. This would suggest that targeting of recombinant protein is efficient with low latency in the cytosol, and that protein processing does not require involvement of ER or other systems examined by the molecular markers used previously.

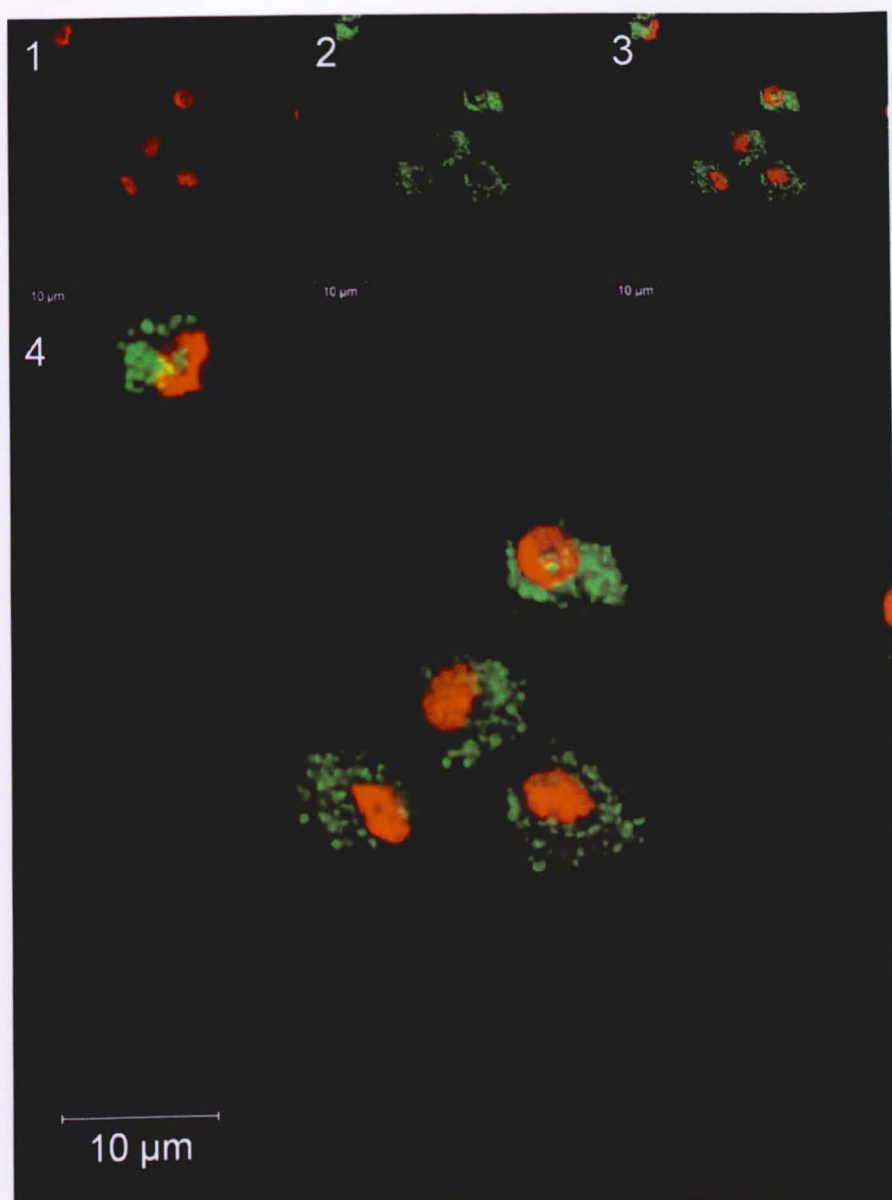


Figure 6.4.3 The localization of the Tim44 *T. vaginalis* homologue

Tim44 transformant *T. vaginalis* C1, were fixed and stained with PI (1) and fluorescently labeled anti-HA antibody (2), combined images (3.4) show localization of the antibody to organelles consistent with the hydrogenosome. Two lasers were used to excite the channels in this image, a 561nm laser was used to excite the PI stain and emission light collected at $>600\text{nm}$, and a 488nm laser to excite the fluorophore coupled secondary antibody whose emission was collected between 493-574nm. The image spans an area $56.41\text{ }\mu\text{m}$ across at a resolution of $0.055\text{ }\mu\text{m/pixel}$.

6.4.2 Localization of Tim44

The inner membrane/matrix translocase associating protein Tim44 was examined and shown in Figure 6.4.3, its distribution would be expected to localize within the inner membrane, however the fluorescently identified protein (channel 2), shows an almost uniform intensity across the organelle, suggesting that Tim44 might exist as a freely distributed matrix protein, as well also

possibly residing with the inner membrane. Reconstruction of organelle surfaces are shown in Figure 6.4.4 and show a similar set of structures to those seen with frataxin.

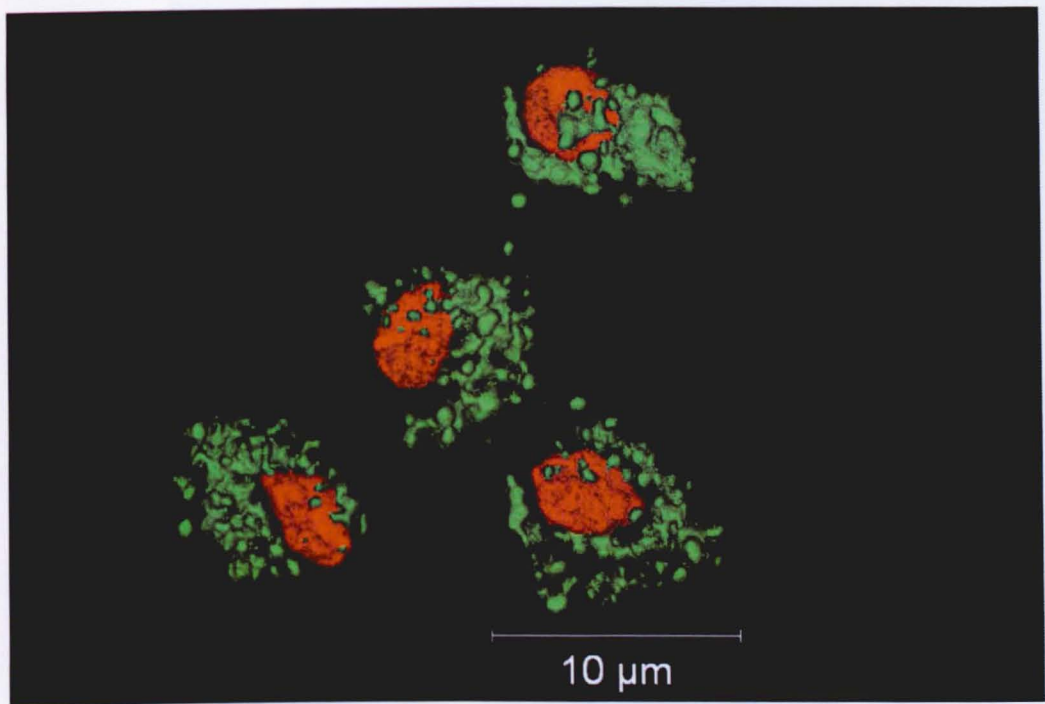


Figure 6.4.4 Reconstruction of hydrogenosomal surfaces in Tim44

The ZEN LE program was used to reprocess the channel composite images seen in Figure 6.4.3 and used to plot iso-surfaces of equal intensity for nuclear and hydrogenosomal compartments seen in this figure. The algorithm used to describe the surface struggles with the organelle crowded regions, but topologically isolated organelles can be seen, with similar size and distribution to the organelles seen previously in frataxin. Additional processing of the original data set in the bioView3D program are seen in Appendix 10.4.6 p257.

The image shows a variety of topologies which could correspond to flattened disks, or cupped shaped compartments – far more so than the previous frataxin images (Figure 6.4.1 p164). If Tim44 has a distribution which transiently associates with the inner membrane, then this would indicate that the matrix compartments in these hydrogenosomes are not uniformly spherical as shown in EM slices in *T. foetus* (Benchimol 2009). The image does clearly indicate a similar kind of distribution to frataxin, and affirms the role and localization of this candidate protein to the hydrogenosome.

6.4.3 Localization of Sam50 homologue Hmp43

Figure 6.4.5 shows the localization of the putative Sam50 homologue Hmp43, an outer membrane β -barrel protein. A composite image (3,4) was produced from a PI channel (1), and the fluorescently identified HA-tagged protein (2).

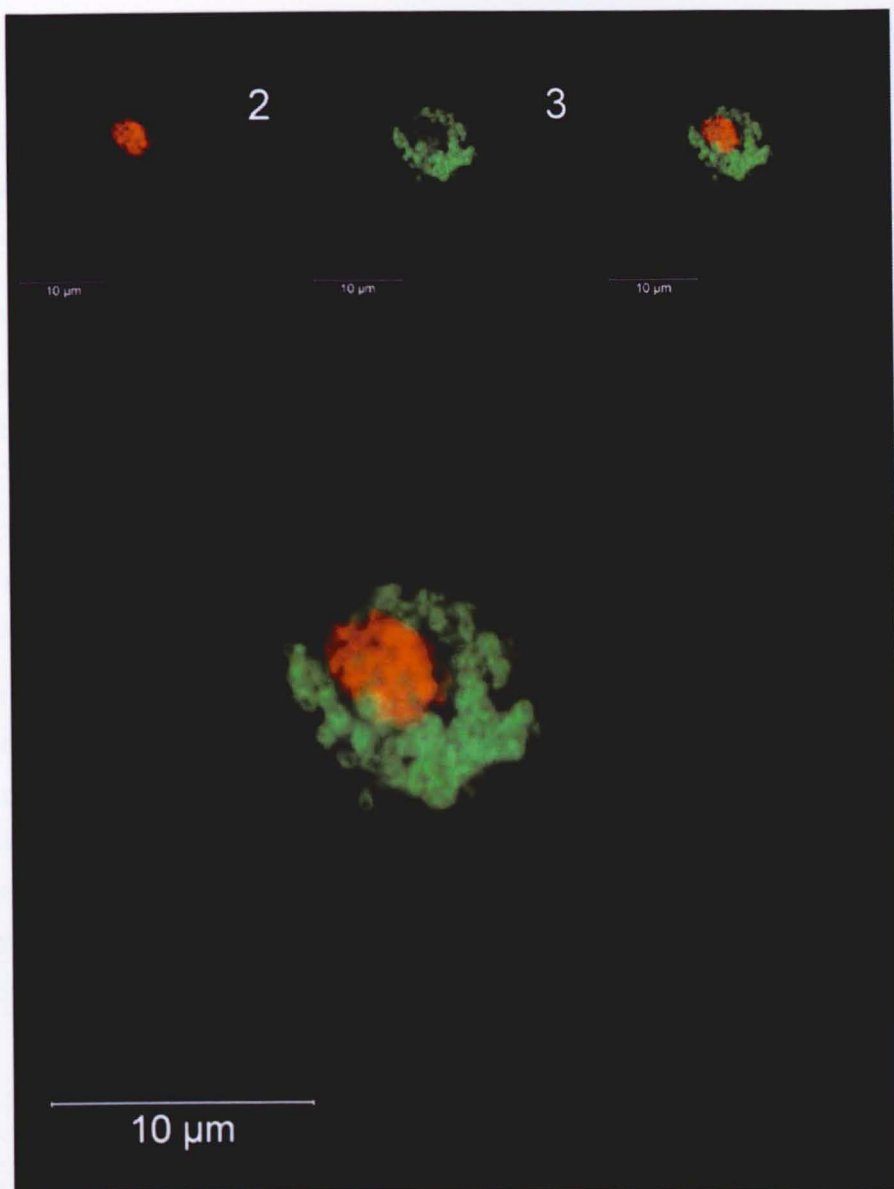


Figure 6.4.5 Distribution of the Sam50 homologue, Hmp43

Hmp43 expressing transformant cells were fixed and stained with PI (1) and Alexa488/anti-HA antibody complex (2). Combined images (3,4) reveal an organelle distribution similar to the frataxin distribution. The single cell shown above was sampled at $0.016 \mu\text{m}/\text{pixel}$, and spans a total area of $32.75 \mu\text{m} \times 33.72 \mu\text{m}$. A 488nm laser was used to excite the Alexa488 and fluorescent signal was collected between 493-555nm, the PI stain was excited using a 561nm and light collected at $>600\text{nm}$.

Distribution of fluorescent signal differs slightly from the previous example seen with frataxin, and the Tim44 homologue. Here there are no uniformly filled organelles, fluorescent signal is always associated with circular or crescent type structures, though organelle density in the shown slice limits the discernment of individual membrane envelopes. The size and distribution of the circular features are however consistent with hydrogenosomes seen previously (Benchimol et al. 1996).

This distribution would indicate that that Hmp43 is does not share the same localization pattern of frataxin or Tim44, and that Hmp43 is membrane bound, or exists within the inter-membrane space. Microscopy alone cannot however differentiate between inter-membrane and membrane protein compartments, and these differences will be explored later in the biochemical analysis section.

6.4.4 Localization of the Hup3 family of proteins, candidate Tom40 homologues

The localization of candidates for the Tom40 translocase, the Hup3s, was imaged as described above for other transformed cells. For clarity only Hups 3b and 3e are shown here, as their distributions are typical of this family, other images relating to the other Hup3 proteins can be seen in the appendices (p258-260). Each image combines a nuclear stained channel using PI (channel 1) and the immunodetection of HA-tagged protein by fluorescent antibody (channel 2). In each case these related proteins consistently show the same pattern of localization.

Fluorescent signal is detected solely at the organelle periphery consistent with the pattern observed with Hmp43. This would indicate specific localization to membrane or inter membrane compartments. The absence of signal outside the hydrogenosome indicates that targeting is efficient and low latency, and does not require transit through other compartments. However without a membrane marker and increased spatial resolution it is difficult to discern the proportion of recombinant protein inserted into the hydrogenosomal membrane, and if there is latent protein in the inter-membrane space.

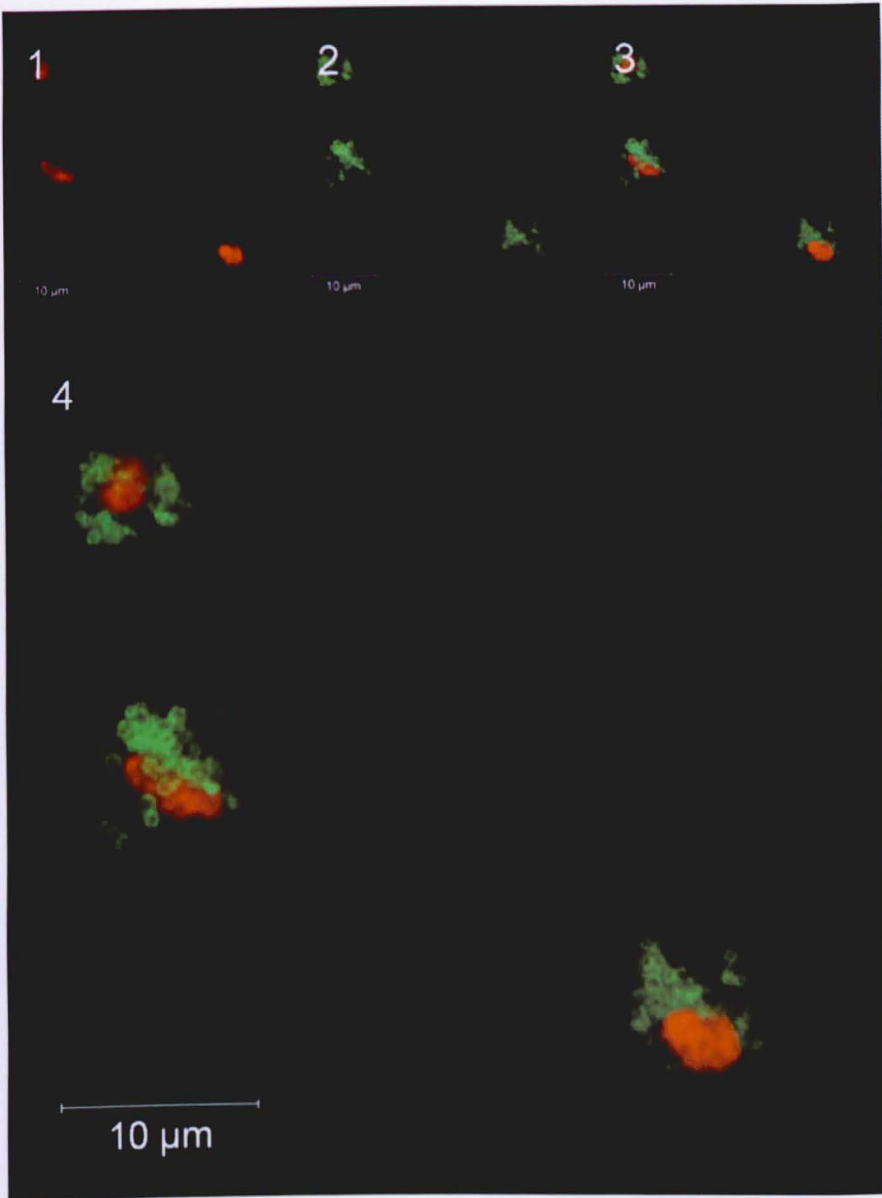


Figure 6.4.6 Intracellular distribution of Hup3b

T. vaginalis C1 cells expressing HA-tagged Hup3B were fixed and stained with PI (1) and fluorescently tagged/anti-HA antibody complex (2). Localization of the HA signal is shown clearly to the hydrogenosomal membranes, showing clear ringlet structure within the confocal image slices. Excitation of the PI channel was with a 561nm laser with light collected >600nm. The Alexa488 secondary antibody was excited using a 488nm laser with fluorescent light collected between 493-555nm. The image resolution was captured at 0.088 µm/pixel.

In contrast to the images obtained for frataxin, Tim44, the organelles visualized for the outer membrane candidates (Hmp43, Hup3 proteins) are typically more regular. This might reflect the model where the organelles are spherical, but the internal topologies are more complex arising from the intrusion of peripheral vesicles. This would fit with prior EM work (Benchimol 2000; Benchimol 2009), and reconcile the differences between the group of inner and outer membrane candidates.

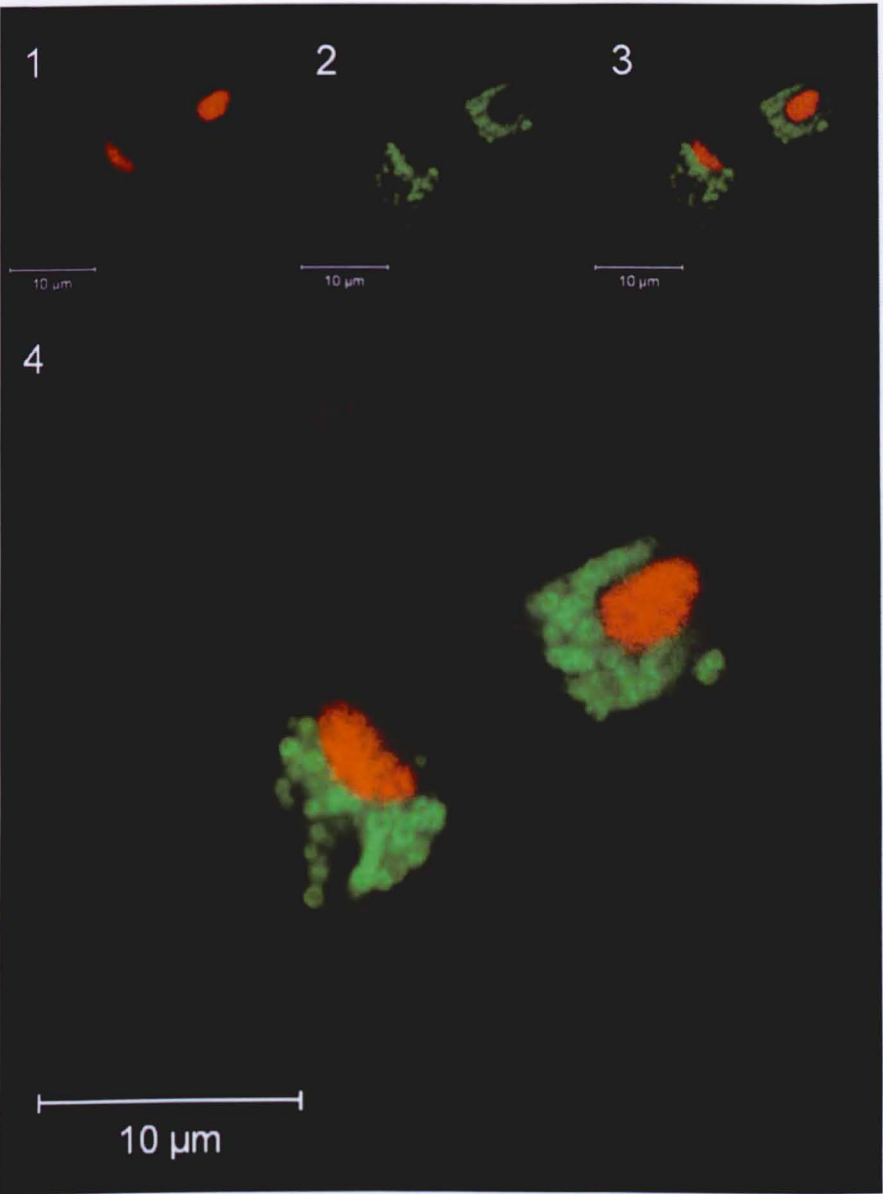


Figure 6.4.7 Intracellular distribution of Hup3e

The hydrogenosomes of the Hup3e transformant strain were visualized by HA immunodetection (2), and their cellular distribution shown against the nuclear marker PI (1). Combined images show a discrete localization of recombinant protein to the hydrogenosomal membranes. Excitation of the PI channel was with a 561nm laser with light collected >600nm. The Alexa488 secondary antibody was excited using a 488nm laser with fluorescent light collected between 493-555nm. The image resolution was captured at 0.066 μm/pixel.

6.5 Further features of the hydrogenosome

The regularity of the hydrogenosome can be seen clearly in Figure 6.5.1 where a high resolution image of Hup3c was obtained. The image shows a sign confocal slice and clearly illustrates the density of the organelles within the cell body. The other clear feature of this image is the relative homogeneity of the shape and size of the hydrogenosome. This can be seen in further detail in the appendices. This would indicate a population of organelles with a regular size distribution, some attempt to analyse the mathematical distribution is made in Appendix 10.4.11. However the limited spatial resolution on the Z-axis imaging constrains estimates on the ability to track individual organelles, and to derive population distributions.

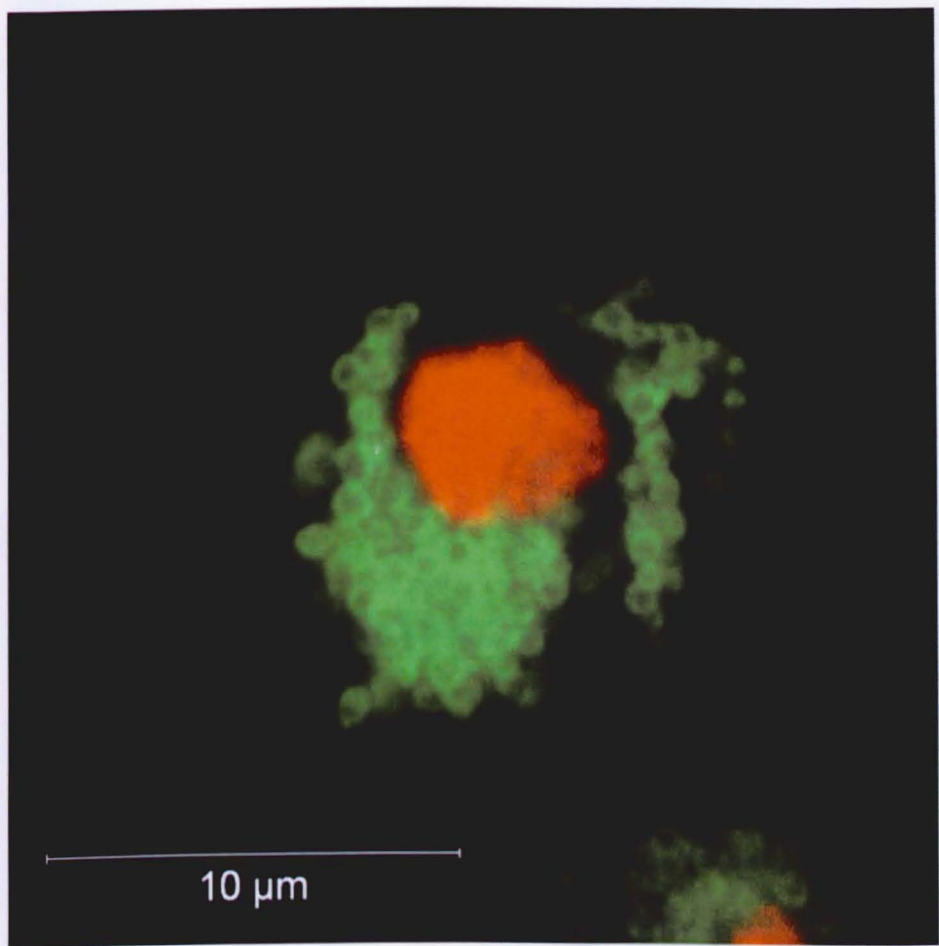


Figure 6.5.1 High resolution microscopy of hydrogenosomes in a Hup3c transformant

A single transformant cell was imaged at ultra high resolution, (0.022 μm/pixel). The nuclear marker, PI, is shown on a red channel, and was excited using a 561nm laser and emitted light collected >600nm. The recombinant protein was detected with an anti-HA primary antibody, which was then detected with a fluorophore coupled secondary antibody. The Alexa488nm fluorophores was excited with a 488nm laser and fluorescent light collected between 493-547nm.

In Summary

This microscopic investigation has provided convincing evidence that the transformants produced from the bio-informatically identified candidates share a common localization which do not co-localise with other membrane compartments. These data would indicate that these related proteins are targeted to the hydrogenosome.

The specificity of HA-tagged recombinant proteins to the hydrogenosome shown in these images indicates that the detectable HA-tag content of collected cellular fractions is representative of hydrogenosomal material. Cell fractionation experiments in the next chapter will build upon these images to show the cellular distribution of HA-tagged recombinant protein.

Whilst the microscopy indicates some evidence for mitochondrial homologues directed to their hydrogenosomal equivalent compartments, the resolution of the images is insufficient to exactly localize the proteins to membranes or inter membrane spaces. To address this biochemical approaches will be used in the next chapter to exactly localize proteins to their functional compartments.

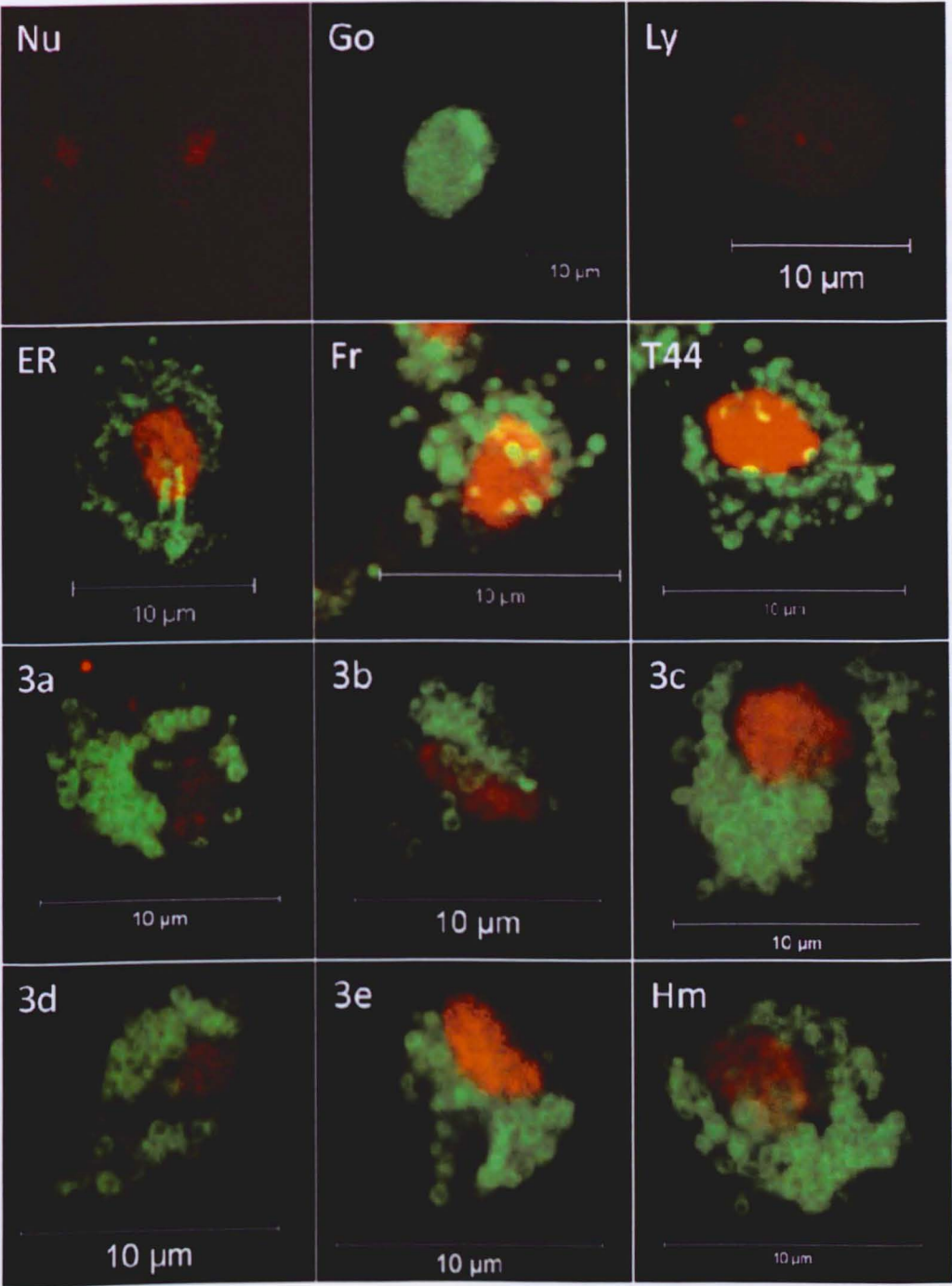


Figure 6.5.2 Localisation patterns of selected cellular markers in *T. vaginalis*

This composite image illustrates the diversity of localisation patterns seen in the microscopy of *T. vaginalis* in this chapter. Distribution patterns for specific cell compartments are shown, Nu; nuclear localisation of Propidium iodide, Go; the ceramide Golgi marker, Ly; lysoTracker, ER; Endoplasmic reticulum anti-BIP antibody. The variety of different recombinant hydrogenosomal proteins are also shown, the matrix iron-sulphur protein frataxin (Fr, green) and Tim44 homologue (T44, green), as well the membrane proteins Hmp43 (Hm, green), and those of the Hup3 family (3a-e, green). Propidium iodide (red) has been used as a nuclear marker in all images using the Alexa488 immunodetection.

7 Biochemical characterization of candidate proteins

7.1 Practical approach to transformant investigation

Bioinformatic and microscopic analysis of candidate proteins have already suggested localization and function of the selected proteins, however the limitations of optical resolution and *in silico* analysis do not provide concrete proof for membrane localization, and will not reveal interaction partners or complexes candidate proteins participate in. This section will see the use of biochemical characterization techniques to corroborate localization data and to inform about native protein function.

Throughout this section practical techniques are developed and built upon to give incremental data concerning the characteristics of labeled proteins in the transformants generated. Firstly localisation studies were developed building upon the organelle extraction strategies in Chapter 4 for import assays, to determine organelle localisation (7.2), but then further elaborated to determine membrane localisation through membrane extraction techniques (7.3). These corroborate predicted localisation evidence seen in microscopy and inferred in bioinformatic investigation.

The isolation strategies developed in Chapter 4 are then employed to provide hydrogenosomes for other molecular biology techniques. With the determination of membrane localisation for transformant proteins (7.3) techniques are used to probe the complexes that the transformant proteins form (7.4, 7.5), and methods are used to purify and isolate these complexes (7.6). Attempts to further characterize complex constituents are briefly discussed, as well as the value the techniques developed in this section with their potential to be employed to advance further aspects of the *T. vaginalis* hydrogenosomal membrane system.

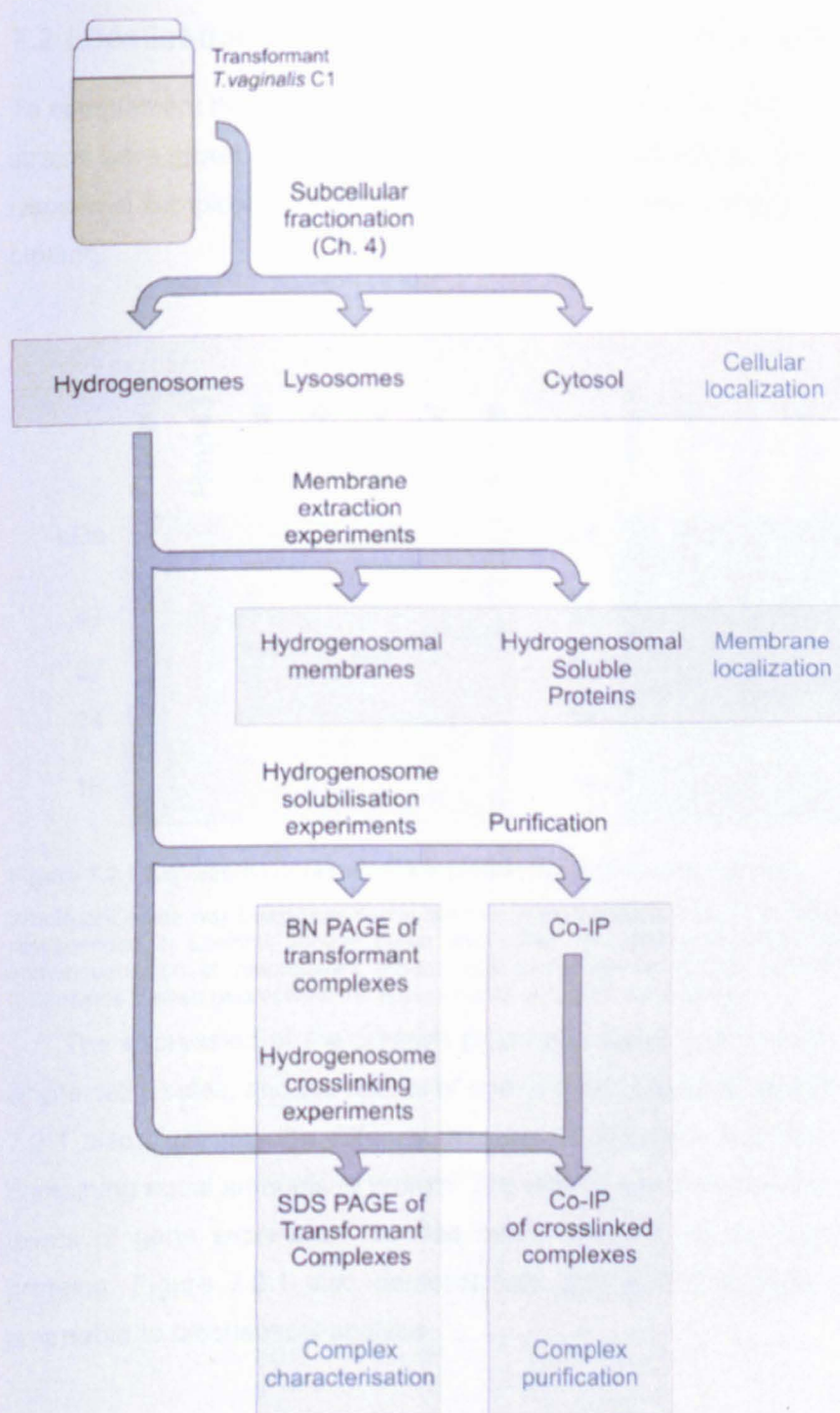


Figure 7.1.1 Practical approaches employed to investigate *T. vaginalis* transformants

Organellar purification strategies developed in Ch 4 are used to produce purified transformant hydrogenosomes, these hydrogenosomes containing HA-tagged proteins are then investigated by a variety of molecular biology techniques, aiming to identify localisation to complement microscopy work, and determine any complexes with which these proteins form. Furthermore strategies to purify recombinant tagged protein and interacting proteins are developed.

7.2 Localisation of candidate proteins in *T. vaginalis* cell fractions

To complement the localisation provided by the microscopy data, cultures of transformant strains were grown, and the collected cells fractioned into cytosol, hydrogenosomal, and lysosomal samples; recombinant protein was then detected by the HA-tag using western blotting.

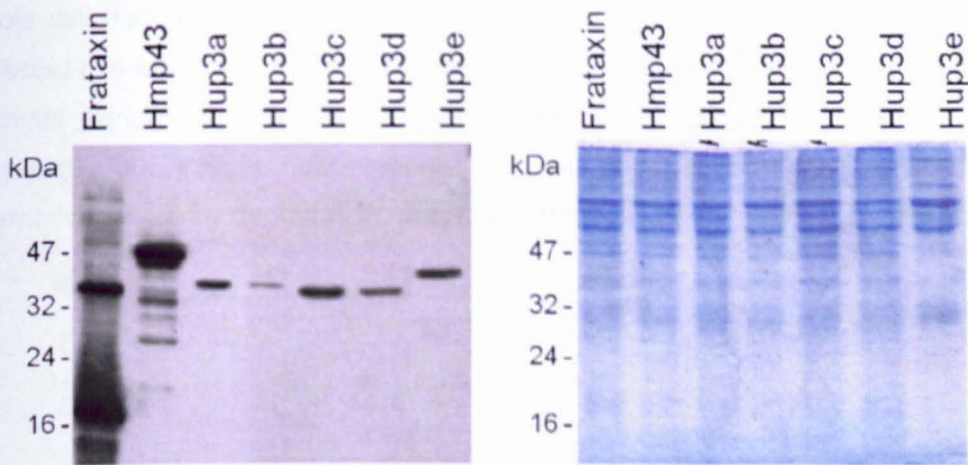


Figure 7.2.1 Expression of recombinant protein in transformant cell lines

Whole cell lysate was taken from dense cultures of transformant cells (2×10^8 cells/ml) and cells pelleted before being resuspended in Laemmli sample buffer and boiled. Samples were then separated on 12% SDS PAGE gels. Immunodetection of recombinant protein was performed on blotted gels using an anti-HA antibody. Parallel Coomassie stained gels verified the approximately equal protein loading.

The expression of the putative β -barrel proteins was first detected in transformant cells in whole cell lysates, and the results of one of these experiments are shown in Figure 7.2.1. Figure 7.2.1 also illustrates the different intensity of detected of protein within samples of cell lysate containing equal amounts of protein. The differences in signal intensity might arise from different levels of gene expression, or else reflect the half-life, or transfer efficiency of the specific proteins. Figure 7.2.1 also demonstrates that all transformants generate a signal intensity amenable to biochemical analysis.

7.2.1 Cellular localisation of recombinant HA-tagged frataxin

To confirm the immunolocalisation analysis in Chapter 6, the cellular localization of candidate proteins was probed following cell fractionation. The HA-frataxin is used here again as a marker for the hydrogenosomal fraction, and results shown in Figure 7.2.2. Samples were collected for

whole cell, cytosol, hydrogenosome and lysosomal fractions. Expression is demonstrated in the whole cell sample, and the greatest concentration of signal is in the hydrogenosomal fraction (H).

Some signal is observed in the lysosomal fraction, this most likely indicates contamination either by vertical mixing or migration of disrupted hydrogenosomes during ultracentrifugation. Alternatively the signal might arise from hydrogenosomal autophagy, but it is difficult to support this claim. Of additional note a double band is observed for protein in the whole cell fraction. As only protein with an intact C-terminal region with intact HA-tag can be detected this would suggest that if there is a real modification to protein mass that there is an N-terminal truncation. This could fit with a maturation of the protein by cleavage of signal sequence, but this is not observed for the hydrogenosomal fraction. The effect could alternatively arise by degradation, or preserved secondary structure in this sample.

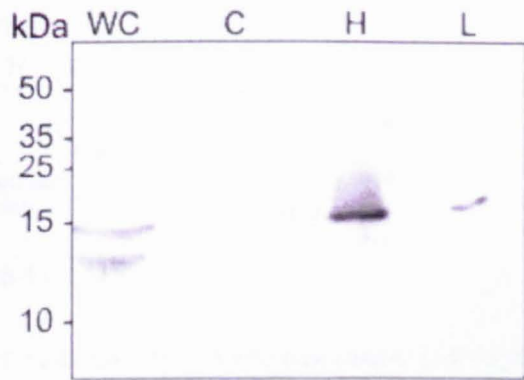


Figure 7.2.2 Cellular localisation of frataxin

The subcellular localization of the hydrogenosomal protein frataxin is shown illustrated above. 2ml of pelleted cells from a late log culture were collected for the Whole-Cell (WC), the equivalent amount of cytosol generated from 2ml of pelleted cells was also loaded (C). Hydrogenosomes produced from the culture (H) were loaded adjacent to a lysosomal sample (L) these samples are of equal total protein, and the hydrogenosomal load is consistent to the whole-cell equivalent (2ml ≈ 50µg). Molecular weight markers were loaded with the samples and their positions are indicated to the right.

The data shown in Figure 7.2.2 would indicate that frataxin is efficiently targeted to the hydrogenosome as no signal is observed in the cytosol fraction. These data would also support that the cell fractionation method is efficient at sorting hydrogenosomes and differentiating them between other organellar compartments (lysosomes), and that the HA-tag does not disrupt the normal targeting of the protein , these data also fit with the microscopy observations of this protein.

7.2.2 Localisation of wild-type and HA recombinant Hmp43

To further support the data provided by the frataxin control, in particular to further demonstrate that tagging candidate proteins does not alter their localization, Hmp43 localisation was tested with an antibody (rabbit) raised against Hmp43 purified from an exogenous *E. coli* expression system (work by S. Dyall and H. Brooks unpublished). The localization of wild-type protein versus the HA recombinant version is shown in Figure 7.2.3 and Figure 7.2.4.

The immunodetection of wild type Hmp43 for the parental strain G3 is illustrated in Figure 7.2.3. The serum shows specificity for Hmp43, when contrasted with the pre-immune serum. Signal is clearly seen for the recombinant *E. coli* protein as well as strong detection in both hydrogenosomal (H) and hydrogenosomal membrane (P) fractions, however signal cannot be seen for the whole cell samples, possibly due to degradation or low endogenous expression.

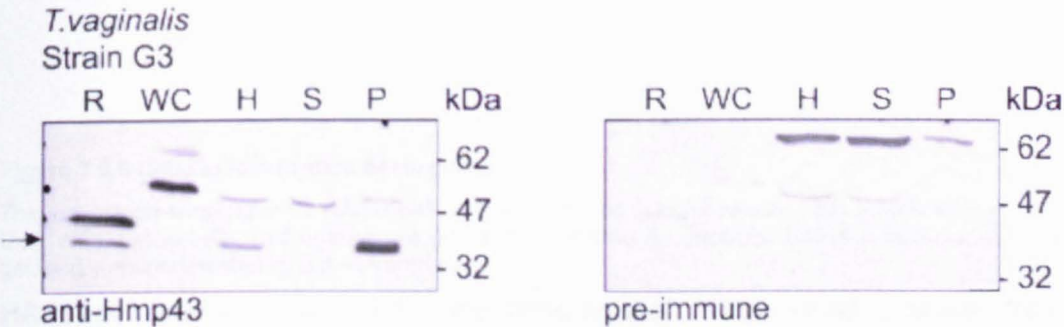


Figure 7.2.3 Localisation of wild type Hmp43 in *T. vaginalis* G3

Recombinant protein (R) generated in *E. coli* from the genomic DNA from *T. vaginalis* G3 was used to raise an antibody against Hmp43, the specificity of the antibody is shown against its pre-immune (right). The antibody was then used to test for Hmp43 in Whole-Cell (WC), and Hydrogenosomal (H) fractions. Additionally fractions obtained by sodium carbonate fractionation for hydrogenosomal soluble proteins (S) and membrane proteins (P) are used to differentiate between membrane and matrix localization.

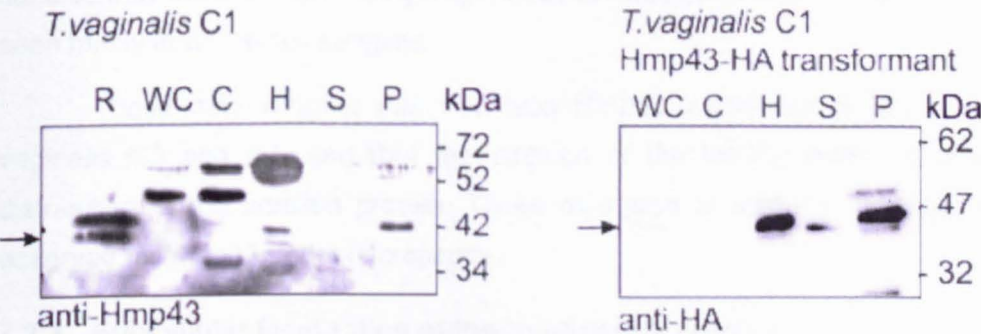


Figure 7.2.4 Localisation of Hmp43 in *T. vaginalis* C1

The anti-Hmp43 antibody was used to show the localization pattern of Hmp43 in *T. vaginalis* C1. *E. coli* recombinant protein (R) was used as a control, and cell fractions corresponding to Whole Cell (WC), Cytosol (C), Hydrogenosomal (H) and Hydrogenosomal soluble (S) and membrane (P) proteins. Antibody detection matches that seen in strain G3.

The cell fractionation pattern is also shown against the HA-Hmp43 C1 recombinant strain, where the HA tagged protein exhibits a similar pattern of localization.

The anti-Hmp43 antibody was also used to detect Hmp43 in the parental C1 strain of *T.vaginalis* (Figure 7.2.4, left) which was found to match that of strain G3 Figure 7.2.3. The localization of tagged HA-Hmp43 was probed in the transformant C1 strain with an anti-HA antibody and is also shown in Figure 7.2.4. The localisation of the tagged protein in the C1 transformant is similar to that of the wild-type Hmp43 detected in the C1 parent strain.

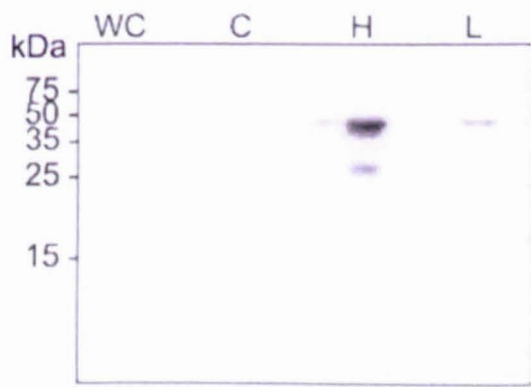


Figure 7.2.5 Cellular localisation of Hmp43

The subcellular localization of HA-Hmp43 (C1) was probed using a cellular fractionation series consisting of Whole Cell (WC), Cytosol (C), Hydrogenosome (H), and Lysosomal (L) fractions. Samples were run on a 12% SDS PAGE gel, and immunodetected by anti-HA antibody.

HA-Hmp43 was also examined in the same type of sample series produced for frataxin, with whole-cell, cytosol, hydrogenosomal, and lysosomal fractions, and is shown in Figure 7.2:5. These samples were prepared for equal loading as for frataxin (Figure 7.2.2), and demonstrates a similar distribution of signal, with most signal present in the hydrogenosomal fraction. Similar to frataxin in Figure 7.2.2 some signal is observed for the lysosomal fraction, though considerably weaker than the hydrogenosomal fraction. In addition recombinant protein can be seen faintly in whole-cell samples.

These data suggest that wild type Hmp43 is directed to the hydrogenosome in *T. vaginalis* G3 and C1, and that the addition of the HA-tag does not alter the sub-cellular distribution of the labelled protein. These data also fit with the hydrogenosomal localisation observed for Hmp43 under microscopy.

7.2.3 Subcellular localization of the Hup3 family proteins.

The localization of the Hup3 family of proteins were determined using the same subcellular fractionation series as frataxin and Hmp43, with Whole-Cell (WC), Cytosol (C), Hydrogenosomal (H) and Lysosomal (L) fractions, and the resulting data can be seen in Figure 7.2.6. All the Hup3

proteins are clearly seen in the hydrogenosomal fraction. Some degree of lysosomal contamination is seen for the Hup3 proteins, though those with highest hydrogenosomal signal intensity also have the highest signal intensity in the lysosomal fractions, this might simply indicate that there is a proportional contamination of the lysosomal fraction with hydrogenosomal material inherent in the organelle purification strategy. The data presented in Figure 7.2.6 corroborates the microscopy data that the Hup3 family are hydrogenosomally targeted proteins.

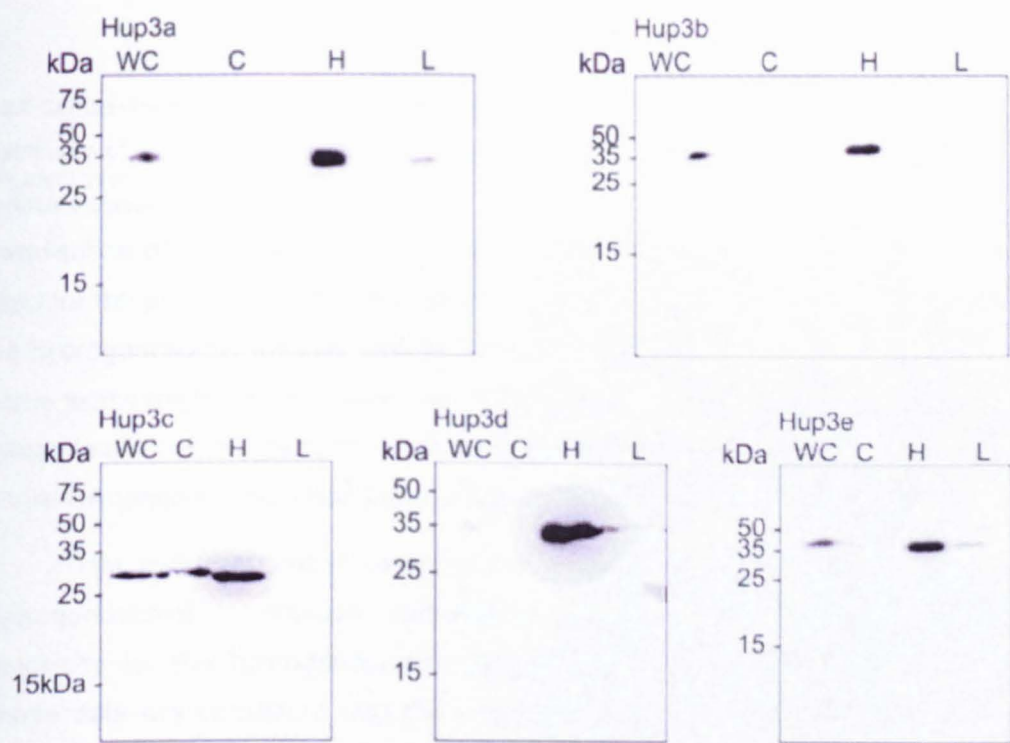
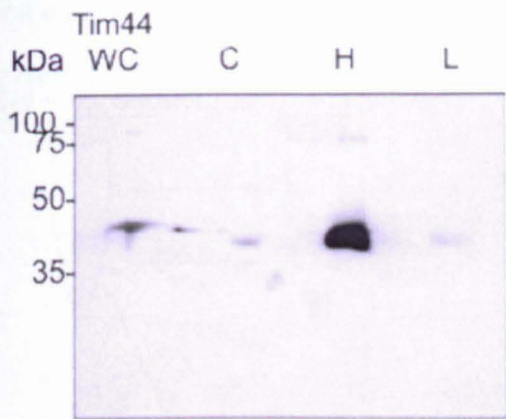


Figure 7.2.6 Cellular localization of the Hup3 family proteins

Cultures of each Hup3 transformant were used to prepare a subcellular fractionation series, Whole Cell (WC) samples, as well Cytosol (C), Hydrogenosomal (H) and Lysosomal (L) fractions were collected. These samples were run on SDS PAGE gels, and blotted, HA-tagged protein was immunodetected by anti-HA antibody.

7.2.4 Subcellular localization of the Tim44 homologue



7.2.7 Cellular localization of the Tim44 homologue

Distribution of HA-tagged Tim44 was investigated in 4 cell fractions, Whole Cell (WC), Cytosol (C), Hydrogenosomal (H) and Lysosomal (L). Samples were run on a 15% SDS PAGE gel, and blotted prior to immunodetection with an anti-HA antibody.

Localisation of the *T. vaginalis* Tim44 homologue was probed using the same fraction series used for the previous candidate proteins. The greatest signal for the HA-Tim44 was detected in the hydrogenosomal fraction, with fainter signal detected in lysosomal and whole-cell fractions. Some signal might also be discerned in the cytosolic fraction, though could represent organellar contamination of this sample, however if genuine, this would represent the only transformant where HA-tagged protein had accumulated to a detectable threshold within this fraction.

The cell fractionation experiments conducted for the frataxin control, as well as the hydrogenosomal translocase candidates, all suggest hydrogenosomal localisation, with specificity for the hydrogenosomal compartment over cytosol or lysosomal compartments. These data are consistent with the localisation patterns observed in microscopy, and would indicate that the candidate proteins are not abundant in the cytosol, or otherwise indirectly targeted to the hydrogenosome.

7.3 Hydrogenosomal membrane localisation of candidate proteins

Confocal microscopy and organellar purification experiments have confirmed the localisation of all candidates to the hydrogenosomal compartment, however do not address the membrane localisation of these proteins. To address sodium carbonate fractionation was developed to further characterise the candidate localisation to either membrane or soluble (matrix, IMS) hydrogenosomal compartments . This technique will build upon the microscopy data by being able to discern whether the protein is membrane or IMS localised. The sodium carbonate

extraction method is mentioned on p75 and illustrated in Figure 7.3.1. The results of the sodium carbonate extraction study are seen in Figure 7.3.2.

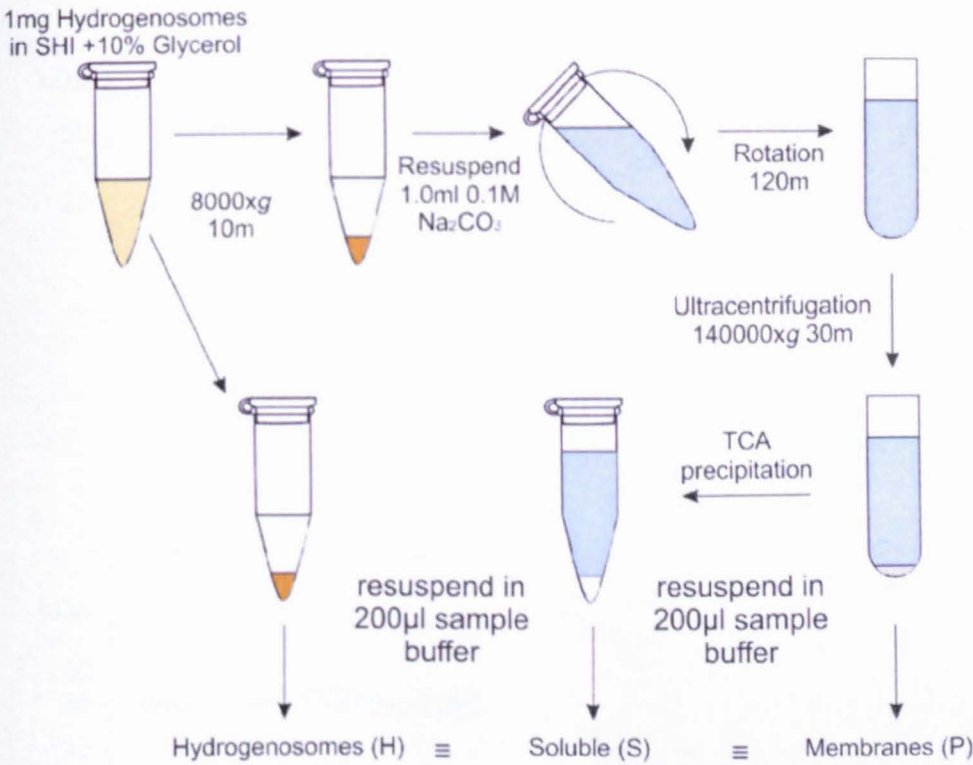


Figure 7.3.1 Derivation of hydrogenosomal membrane and soluble protein samples

Purified hydrogenosomes were used to generate samples for soluble hydrogenosomal proteins, and hydrogenosomal membrane proteins. Sodium carbonate was used to disrupt organelles, and the membrane rafts were collected by ultracentrifugation. Soluble proteins in the supernatant were precipitated by TCA precipitation. This method is described fully on p74. The equivalence between samples is clearly shown in this diagram.

The Sam50 homologue, Hmp43 is seen to unambiguously target to the hydrogenosomal membranes, a pattern which is also observed for the Hup3 proteins. The clarity of this discrete compartmentalisation is blurred a little for Hup3e due to abundant expression of protein in this strain, however the same pattern for preference to the membrane compartment is observed.

These data would suggest that all candidate β -barrels are membrane inserted. These data would indicate that the circular structures observed under microscopy (summarised p175) correspond to immunodetection of recombinant protein within the hydrogenosomal membranes.

Tim44 and frataxin have been observed in previous data to be hydrogenosomally targeted, but with distinct distribution viewed under microscopy. The fractions generated by sodium carbonate fractionation generate different results to those seen for the membrane β -barrel proteins. Frataxin shows near equal distribution to all fractions with a slight bias for the

soluble compartment. The frataxin transformant is the only protein which is strongly present in the soluble fraction, this data fits with data previously collected about the function of this protein.

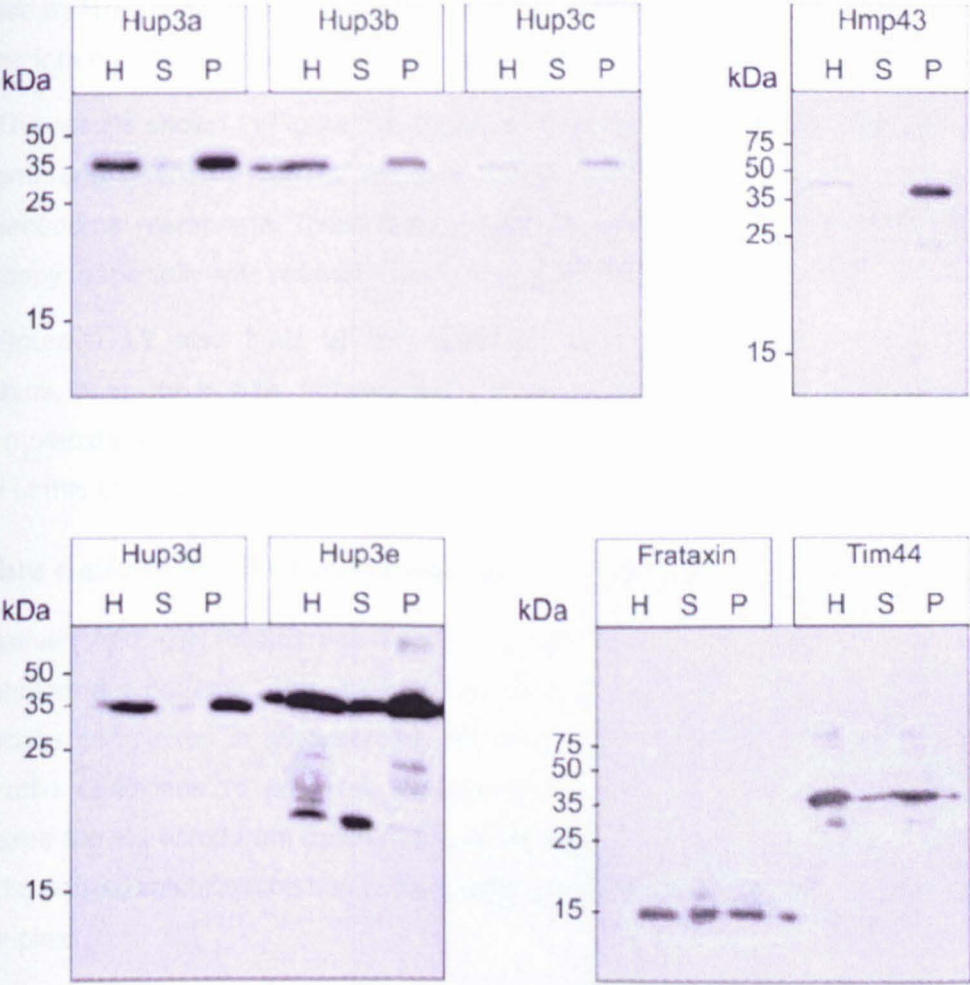


Figure 7.3.2 localization of candidate proteins to the hydrogenosomal membrane

Whilst organelle fractionation identifies that all identified hydrogenosomal translocase candidates target specifically to the hydrogenosome (H), sodium carbonate membrane extraction was performed to determine whether candidate proteins reside within hydrogenosomal membranes (P) or within the soluble compartments of the hydrogenosome (S). Samples in this figure were loaded equally, with 100 µg of hydrogenosomes loaded (H) and the equivalent in the soluble and membrane extracts.

Tim44 shows a different distribution to frataxin and a greater quantity of the protein is detected in the membrane fraction. Tim44 is proposed to be only peripherally associating with the Tim23 complex, however the *T. vaginalis* homologue still seems associated primarily with the membrane fraction after sodium carbonate fractionation, though a subpopulation is observed in the soluble fraction. These data partly contrast with some of the images obtained from microscopy which suggest a more uniform luminal distribution for this protein- though some circular species were also observed (p167). These data might suggest that there is a

subpopulation of soluble protein within the hydrogenosomal matrix, perhaps as a processing intermediate. This accumulation could reflect disruption by the HA-tag, or from saturation of the insertase by HA-Tim44 over expression, however the HA-tag does not preclude the protein from inserting into the membrane, as the majority of the protein appears to have this localisation.

The results shown in Figure 7.3.2 support the notion that the Hmp43/Hup3 transformants are membrane inserted β -barrels, and that Tim44 might have a tightly associated role with the hydrogenosomal membrane. These data are also consistent with the fine details observed in the microscopy, especially with respect to the circular structures observed for the hydrogenosome.

Figure 7.3.2 also hints at the existence of higher molecular weight SDS resistant complexes, seen for Hup3e, frataxin, and Tim44. To follow from the localisation studies, the higher molecular weight complexes to which the transformants belong are examined in the latter portion of this chapter.

7.4 Blue native PAGE of outer membrane hydrogenosomal candidates

Blue Native PAGE (BN PAGE) was chosen as a method to investigate the complexes formed by the transformant proteins. This method has been used successfully to characterise different translocase complexes in mitochondria (W. Meisinger et al. 2001). BN PAGE is performed under mild conditions to preserve complexes in their native conformation. These protein complexes are extracted from the membrane but they are not additionally stabilised by chemical methods, and so solubilisation has to be gentle enough not to disrupt the proteins, or destabilise the complex.

For this investigation three detergents were found to be particularly effective at solubilising the β -barrel candidates, n-dodecyl beta-D-maltoside (DDM), Digitonin, and Triton X-100. Thawed transformant hydrogenosomes were pelleted and resuspended in solubilisation buffers with these detergents (Methods 2.9.1). Hydrogenosomes were allowed to incubate with rotation, before the detergent lysates were cleared by ultracentrifugation. The supernatants of these solubilisation reactions would contain detergent solubilised membrane complexes, including transformant protein complexes.

Resolution of these complexes was performed under non-denaturing BN PAGE (Methods, p71) and run alongside a BN PAGE marker series, composed of protein complexes (HMW Calibration Kit for Native electrophoresis, GE Healthcare) which would provide native molecular weight information and indicate whether electrophoretic conditions were degrading loaded complexes. Proteins within the resolved gels were then denatured with SDS prior to

western blotting where transformant protein was immunodetected by anti-HA antibody. Whilst the complexes were destroyed after BN-PAGE the positional information of the complexes are preserved in the blots. The results of this experiment can be seen in Figure 7.4.1.

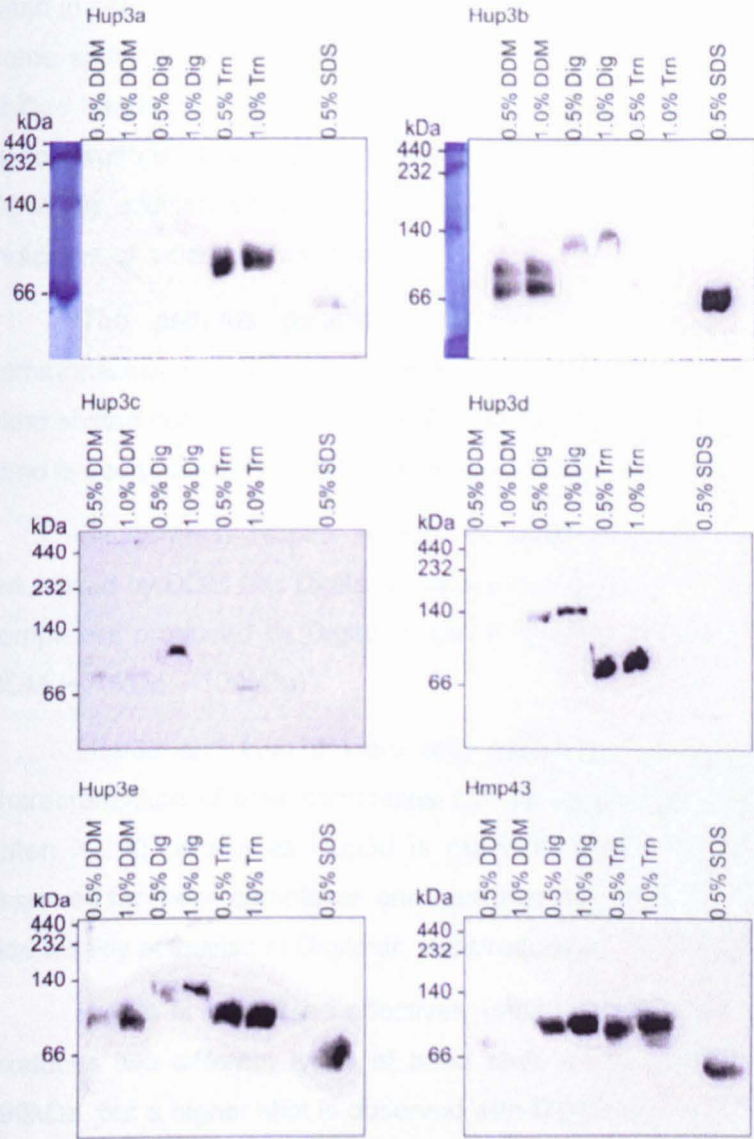


Figure 7.4.1 Blue native PAGE of Hmp43 and Hup3 family proteins under different detergent conditions

To examine the functionality of the candidate proteins BN PAGE was used to probe for any complexes which the proteins might form within the membrane. These complexes are further characterized by their behavior in different detergent conditions. The figure explores the use of four detergents, n-dodecyl beta-D-maltoside (DDM), Digitonin, and Triton X-100, the denaturing detergent SDS, the non-detergent constituents of the solubilisation buffer are described in the methods section. Native molecular weight markers were used (named in text), and their positions are denoted to the left, these were detected visually on staining (sample strips shown top).

The results of the BN PAGE show diversity between individual candidate proteins, this is not unexpected as the sequence diversity of Hup3 is broad, and their reactions with specific detergents is ultimately determined by exposed residues. Hmp43 is seen to generate a shifted band in series treated with Digitonin and Triton X-100 versus the SDS denatured control. Whilst some shift can be anticipated from the conformational shape of the fully folded protein, the shifted bands for Hmp43 are observed around 90kDa. These complexes are observed at both concentrations of detergent showing good solubilisation in Digitonin and Triton X-100, but not DDM. In addition some band intensity is shifted downward in 1.0% Triton X-100, perhaps indicative of a destabilised complex.

The patterns generated for the Hup3 proteins are complex, and show both commonalities and distinct differences in ease of solubility and band shift. Hup3a produces a band shifted complex (~90kDa) with Triton X-100 treatment, but not with DDM or Digitonin. This band is seen relatively shifted with respect to the SDS control.

By contrast Hup3b is not efficiently solubilised in Triton X-100, but is effectively solubilised by DDM and Digitonin, whereupon it produces shifted complexes. Unlike Hmp43 the complexes produced by Digitonin are different in weight (~120kDa) to the shifts produced by DDM (~70kDa, ~100kDa).

Hup3c and Hup3d were only sparingly soluble in the tested detergents but some characterization of their complexes can be made. Hup3c appears more soluble in DDM than Triton X-100, whereas Hup3d is markedly more soluble in Triton X-100. The band shifts observed for these complexes are approximately equivalent (~80kDa). These two transformants also weakly solubilise in Digitonin and produce a ~120kDa (Hup3c), ~140kDa (Hup3d).

Hup3e is solubilised effectively under all three detergent conditions. Like other Hup3s, it produces two different types of band shift, under DDM and Triton X-100, a shift is seen at ~90kDa, but a higher shift is observed with Digitonin ~120kDa.

The data shown in Figure 7.4.1 suggest that the β -barrel candidates may form multimeric complexes, however there is only limited information regarding their composition, and limited resolution of their size. These complexes are also substantially smaller than BN PAGE resolved TOM complexes.

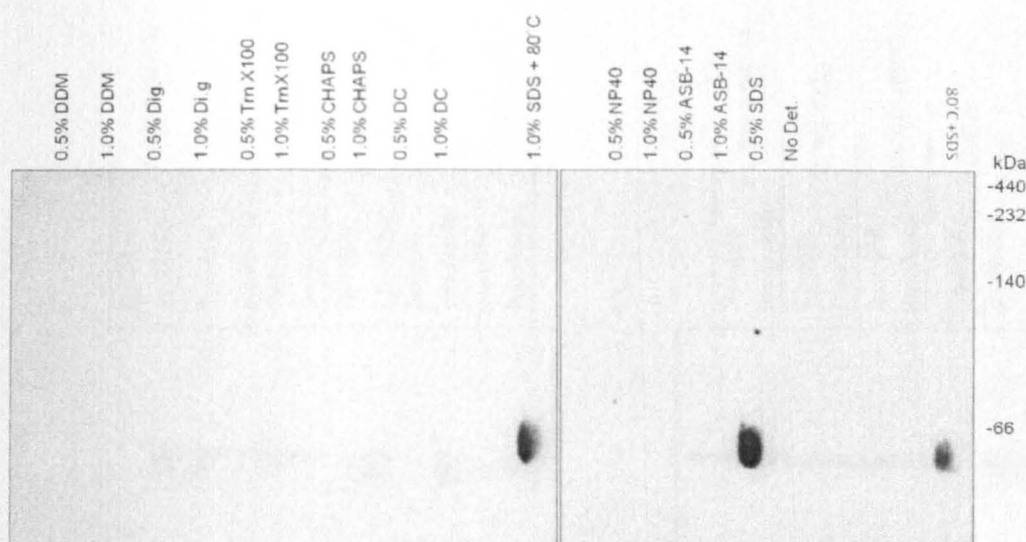


Figure 7.4.2 BN PAGE of solubilised fractions from transformant Tim44 hydrogenosomes

Samples from the solubilisation trial were assayed on a Blue Native PAGE gel, SDS the most efficient detergent in the solubilisation study returns a strong signal corresponding to the monomeric Tim44 protein, however no discernable signal can be seen for any other detergent series.

The Tim44 homologue was also examined using BN PAGE, however there appeared to be significant obstacles to its efficient solubilisation. Attempts to solubilise Tim44 in a more diverse series of detergents is illustrated in Figure 7.4.3, with duplicates of the solubilised fractions shown in a BN PAGE in Figure 7.4.2. Deoxycholine appeared to improve solubility, but was mostly likely as a result of incomplete clearing of detergent lysates. Some solubility is observed for ASB-14-4 and NP-40, however no signal was observed in subsequent BN PAGE. Efficient detection of Tim44 was observed with the denaturing detergent SDS, however bands recovered for the solubilised protein did not indicate any higher molecular weight species.

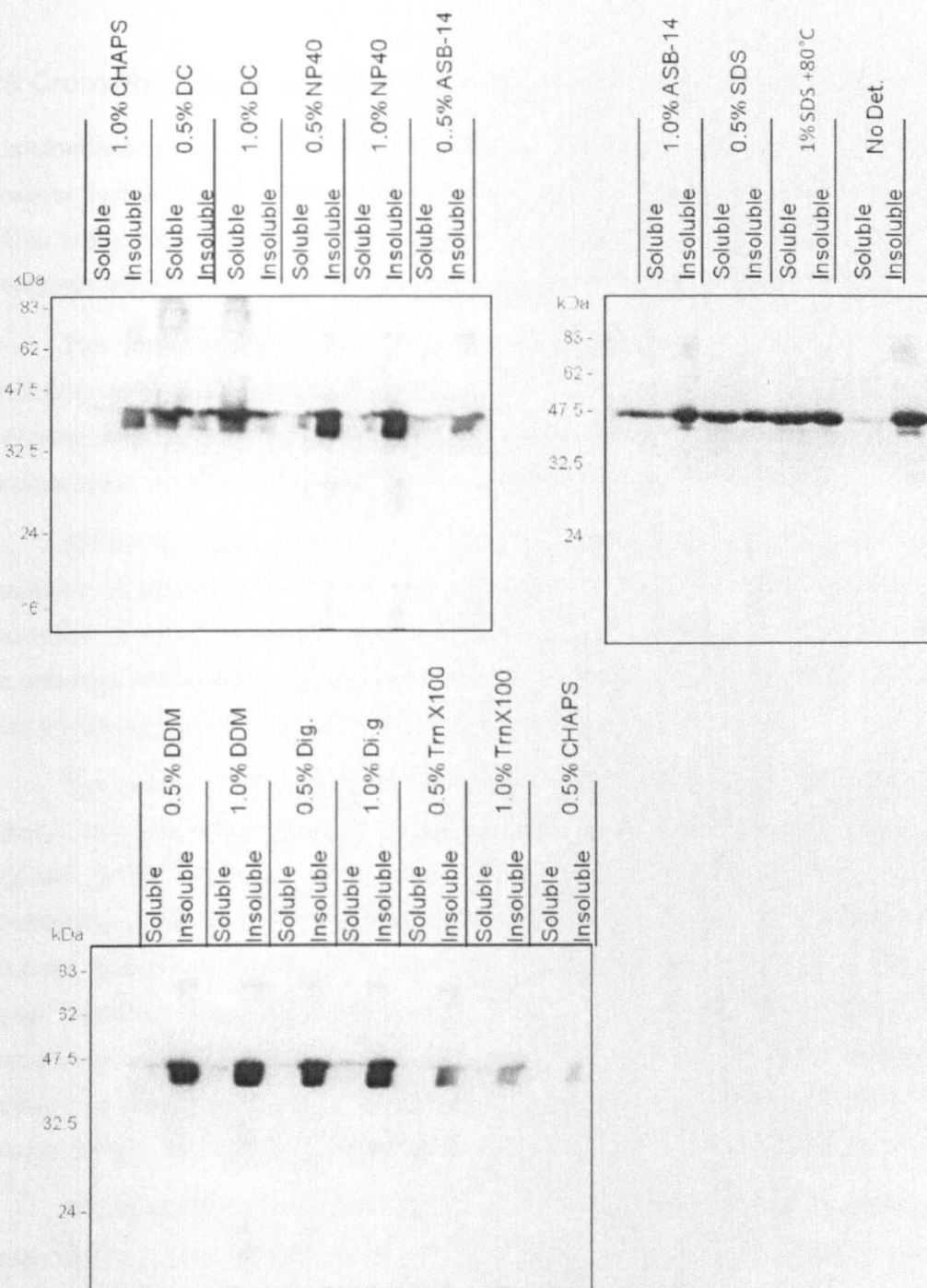


Figure 7.4.3 Solubilisation profile for *T. vaginalis* Tim44 under different detergent conditions

The solubilisation of the Tim44 transformant was assayed on SDS PAGE gels. Different solubilisation solutions were used to try and solubilise Tim44 from hydrogenosomal material. Detergents suited for solubilising the Hup3 and Hmp43 proteins (DDM, Digitonin, Triton X-100) seem to only very weakly solubilise Tim44, whilst the majority of the protein remains in insoluble material. Other detergents were explored, Deoxycholine (DC) generated a turbid lysate which could not be adequately cleared, leading to signal in the soluble fraction, whilst other detergents (CHAPS, NP-40, ASB-14) did not prove any more effective. The denaturing detergent SDS did however efficiently solubilise Tim44, both at room temperature and after heat treatment (80°C)

7.5 Crosslinking of hydrogenosomal translocases candidates

Transformants probed by BN PAGE seemed to indicate the presence of membrane complexes, however their abundance and efficiency of extraction provided obstacles to characterisation. To further study these complexes the covalent crosslinking technique originally employed with the preprotein import assays was redeveloped for probing transformant complexes.

This work would reuse the crosslinking agents DSP and MBS as they could both be employed under the same conditions, but offer a choice of different crosslink bridge spans. The optimisation of crosslinking for transformant hydrogenosomes was an important exercise as the same process would be employed for many subsequent analytical methods.

Whilst the import reactions in Chapter 4 required the optimisation to recover the small quantities of traceable precursor, the transformant hydrogenosomes would have a sizable population of transformant proteins inserted into their membranes. Experiments to determine the optimum concentration of crosslinker to recover recombinant protein from hydrogenosomes were performed with the Hmp43 transformant, shown in Figure 7.5.1.

Strong signal was observed for all test concentrations with both cross-linkers (0.1, 0.5, 1.0mM), however some insolubility was observed at the higher concentrations, indicated by a reduced protein intensity on Ponceau S staining. This effect is expected when excess crosslinking leads to the formation of insoluble protein aggregates. To balance the effects of solubility versus proportion of protein crosslinked the 0.5mM concentration was chosen, though signal intensity was favourably robust across all samples. The effects of crosslinking concentration did not seem to vary between crosslinkers, with MBS possibly showing less recovery at low concentrations. Both crosslinkers generated molecular species which were not seen in the control reactions. The fine structure of these complexes is discussed later.

The reversibility of the DSP crosslinker was also optimised, as chemical cleavage of the cross linking agent would be necessary to isolate protein components from membrane complexes. Treatment of crosslinked hydrogenosomal protein was reversed with incubation with 5mM DTT at 37°C for 15min, (Figure 7.5.2). DTT incubated samples lost most of the high molecular weight complexes from crosslinking and resembled the control reactions. DTT reversal appears effective at all concentrations of crosslinker, which did not influence choice for crosslinker concentration for DSP. DTT treatment also appears to increase soluble protein

content within the reaction volume, with improved protein recovery seen on Ponceau S staining with respect to the crosslinked samples.

The optimisation of crosslinking most critically produced transformant protein crosslinked species. The abundance of these is especially clear with the Hmp43 (Figure 7.5.1). Crosslinking generated a spectrum of species, with the two different agents capturing different complexes. These species are observed at molecular weights higher than the original transformant protein and so represent multimerically linked proteins of varying sizes. The unique species generated by crosslinking range from close to the unlinked control protein, to in excess of 220kDa.

Crosslinking reactions were performed with whole hydrogenosomes and so all possible complexes of the transformant proteins within the hydrogenosome are captured. These diverse complexes would include the final configurations that these proteins adopt, but also assembly and targeting intermediates. Complexes composed of transformant protein with chaperone cytosolic factors, or translation machinery would not be present in this population as no cytosol was incubated in the crosslinking reaction.

Signal intensity in the observed species is highly variable, and reflects both the abundance and the chemical availability to the crosslinking agents. It is thus not possible to infer that the most intense signals represent the most abundant species, but rather represent species which are abundant, and are able to be crosslinked.

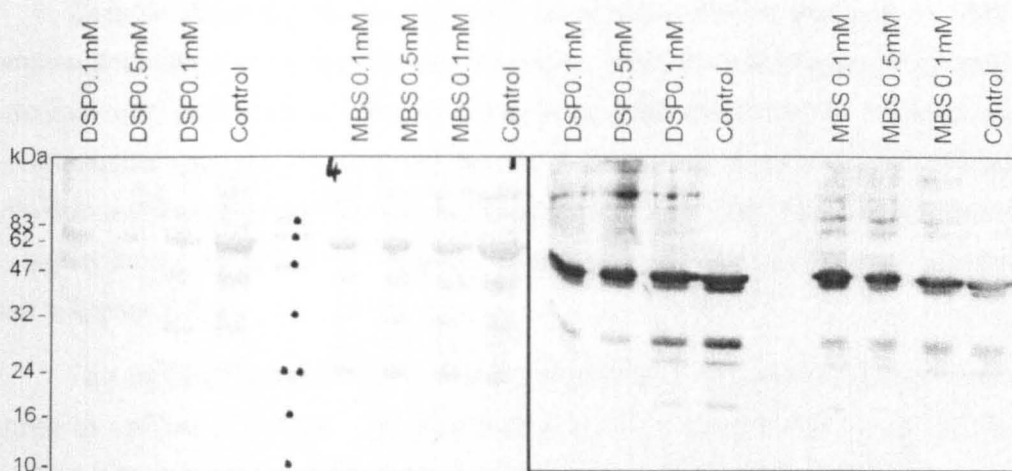


Figure 7.5.1 Optimization of the crosslinkers MBS and DSP on HA-Hmp43

Purified hydrogenosomes from the Hmp43 transformant were subjected to crosslinking as described previously (0 p75). In this assay two crosslinkers, MBS and DSP were tested at three concentrations (0.1, 0.5, 1.0mM). The effect of the crosslinkers on protein solubility is shown on the PonceauS stained blot (left), equal amounts of protein were loaded, however with increasing concentration of crosslinker less protein solubilises for electrophoresis. This blot was then western blotted (right) with an Anti-HA antibody. Hmp43 exhibits a variety of crosslinked complexes seen at higher molecular, these unique to the cross linked reactions versus the control. Increasing concentration of crosslinker, seems to reduce the crosslinked species signal on the blot, likely due to increasing insolubility.

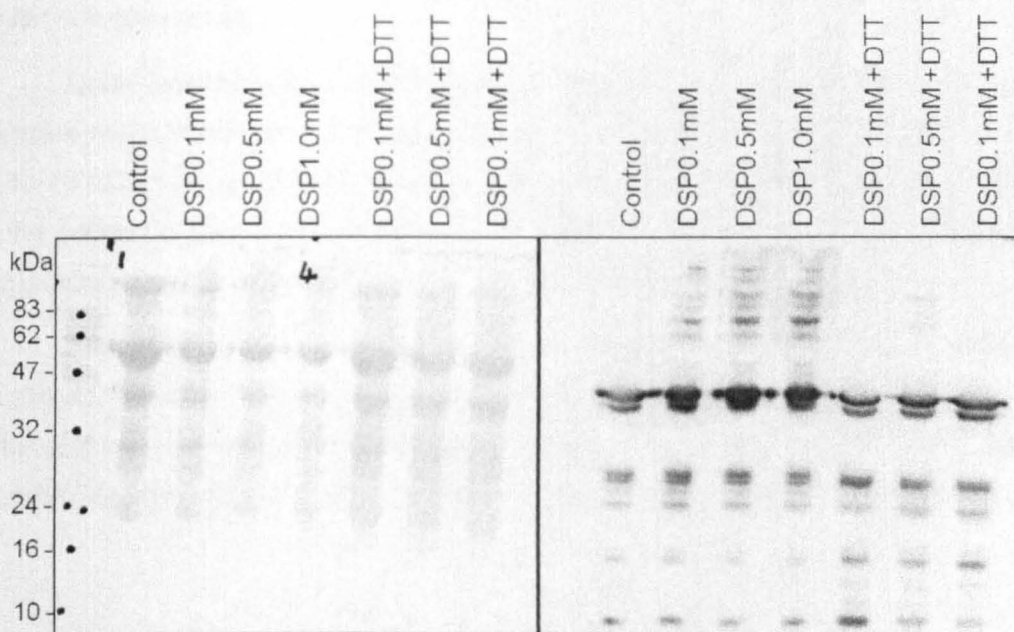


Figure 7.5.2 Reversal of DSP crosslinking in Hmp43 transformant hydrogenosomes

Reversible cleavage of DSP crosslinking was investigated with the Hmp43 transformant. Equal samples treated with DSP crosslinker at 0.1, 0.5, 1.0mM, were loaded alongside duplicate samples treated with DTT to resolve crosslinks. The Ponceau-S stained blot (left) shows that DTT incubation reduces the increasing solubility associated with DSP crosslinking. The effects of DSP cleavage are shown in HA-antibody western blot (right), where it is clear that DTT treatment reduces the majority of crosslinked protein.

Despite these limitations, the trials to optimise the crosslinking in Hmp43 revealed a complex population of covalently linked species. With the establishment of robust crosslinking technique with optimised conditions these were then expanded to improve resolution of the supramolecular species. The first refinement of this technique would repeat the same conditions but within a different gel system, where lower percentage gels were used to resolve masses for the higher molecular weight complexes. This was employed for the other transformants, and is seen in Figure 7.5.3.

The patterns generated for these transformants show remarkable diversity and some degree of unique character. The Hup3 proteins do exhibit some similar crosslinking effects. Matched concentrations of hydrogenosomes yield different quantities of each transformant, perhaps indicating different expression levels, this effect is pronounced in Hup3e, however the remaining Hup3s have comparable intensities.

Amongst the comparable Hup3s the intensity of crosslinking is variable, despite fixed concentration of crosslinker to hydrogenosomal material. This might indicate different availability of suitably placed crosslinkable residues in transformant protein complexes, or abundance of multimeric complexes.

Commonalities also exist between the Hup3s, particularly a 65kDa-80kDa complex. This complex might represent the shifted species seen for the Hup3s in BN PAGE, similarly a faint ~140-160kDa species can be observed (particularly clearly in Hup3c,d,e) and might correspond to the second higher weight species in the BN data. These data support the existence of two complexes, determined both by BN PAGE and crosslinking/SDS PAGE.

Resolution of the complex to sharper detail is hindered by the additional mass of crosslinker residue, whose mass contribution would vary between sparse and complete saturation of susceptible residues (in the order of few kDa). However the abundance and crosslinking characteristics of some species do allow more precise determination of mass.

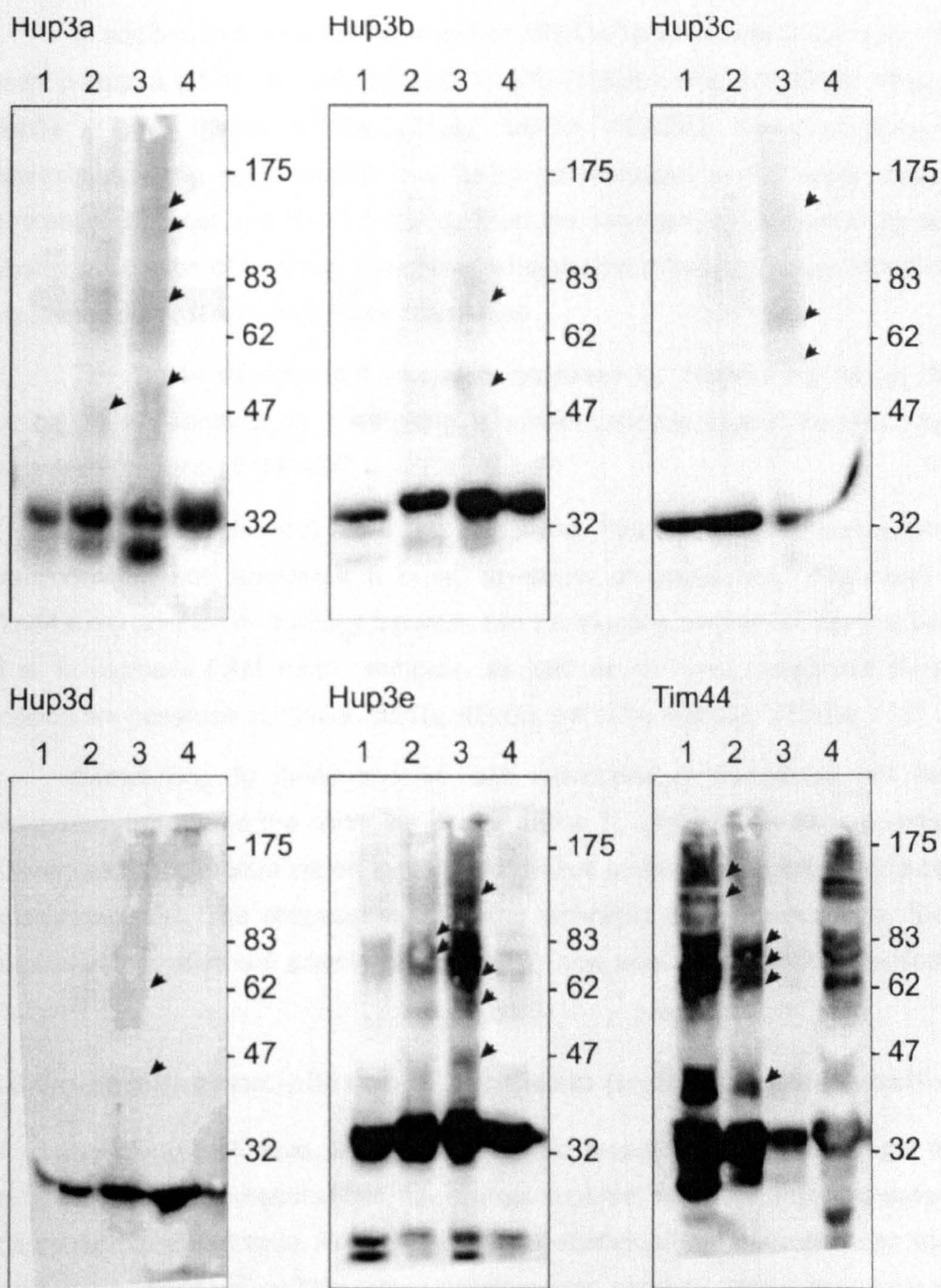


Figure 7.5.3 Crosslinking of candidate hydrogenosomal translocase proteins

To independently examine the protein associations of the candidate translocase proteins, purified hydrogenosomes were subjected to chemical crosslinking with two different crosslinking agents, 3-Maleimidobenzoyl-N-hydroxysuccinimide ester (MBS, 0.5mM, lane 2) and the cleavable crosslinker Dithiobis (succinimidyl propionate) (DSP, 0.5mM, lanes 3,4), which is also shown cleaved with DTT (0.5mM DSP, lane 4). Proteins were run on 8% SDS PAGE gels, and immunodetected by western blot with a anti-HA primary antibody. Selected crosslinked species have been indicated (arrows) and are tabulated later on p203.

In addition to the 60-80kDa, and 140-160kDa species, additional resolved bands can be seen in Hup3a (at 47, 50, 140, 150kDa), Hup3b (55kDa), Hup3c (57kDa), Hup3d (47kDa), and Hup3e (47kDa, 60kDa, 65kDa, 75kDa, 85kDa, 150kDa). Many of these bands are of approximately the same weight, and might be attributed to the slight mass and migration differences between the Hup3 proteins. Thus the apparent diversity of weights might reflect a smaller population of common complexes whereby the different masses of the individual Hup3s are responsible for apparent mass differences.

The Tim44 transformant was also examined by crosslinking, whilst the protein does exhibit strange behavior on crosslinking, it is more tractable to analyze the complexes of Tim44 via crosslinking than BN PAGE.

Tim44 is well recovered at the same hydrogenosomal concentrations as other transformants, and generates a broad spectrum of complexes. This result would support Tim44's key position as a bridge between two translocons, and would capture both the elements of a *T. vaginalis* PAM motor complex, as well as an inner membrane translocon. Distinct species are observed at 45kDa, 60kDa, 65kDa, 140kDa, 150kDa, 175kDa.

Comparison to these species with mitochondrial complexes will be made in the discussion, but due to the divergent nature of the *T. vaginalis* hydrogenosome, and potential differences to the protein import system it is difficult to determine complexes based upon purely crosslinking data. The optimized crosslinking technique does however offer the opportunity to manipulate transformant protein complexes in new way which will be developed later in this section.

7.6 Co-immunoprecipitation of candidate proteins and interaction partners

The crosslinking technique is potent enough to reveal complexes amongst the complicated mixture of proteins present within the hydrogenosome, however it cannot separate proteins of interest from the multitude. To facilitate further studies of the intermolecular interactions of the transformant proteins, co-immunoprecipitation was used to extract transformant protein from hydrogenosomal lysates.

Co-immunoprecipitation builds upon the solubility assays explored by BN PAGE, complexes solubilised by detergents are extracted from the hydrogenosomal lysate by anti-HA antibody. Immunoglobulins are then precipitated with their bound antigens by incubation protein-A Sepharose. Finally these immunocomplexed proteins can be eluted from the resin. By this

method transformant protein complexes identified in the BN page can be purified, so that the eluate only contains proteins which interact through HA-tagged transformant proteins.

This investigation was first tested using Triton X-100 to solubilise transformant protein complexes. Complexes are observed for all putative outer membrane proteins albeit with different efficacy in BN PAGE, co-immunoprecipitation in the same solubilization conditions should recover the same complexes from the hydrogenosomal lysate. These complexes are then denatured and run on SDS PAGE.

Electrophoretically separated proteins are visualised both by sensitive Coomassie staining, and by western blotting (against anti-HA), this data is presented in Figure 7.6.1.

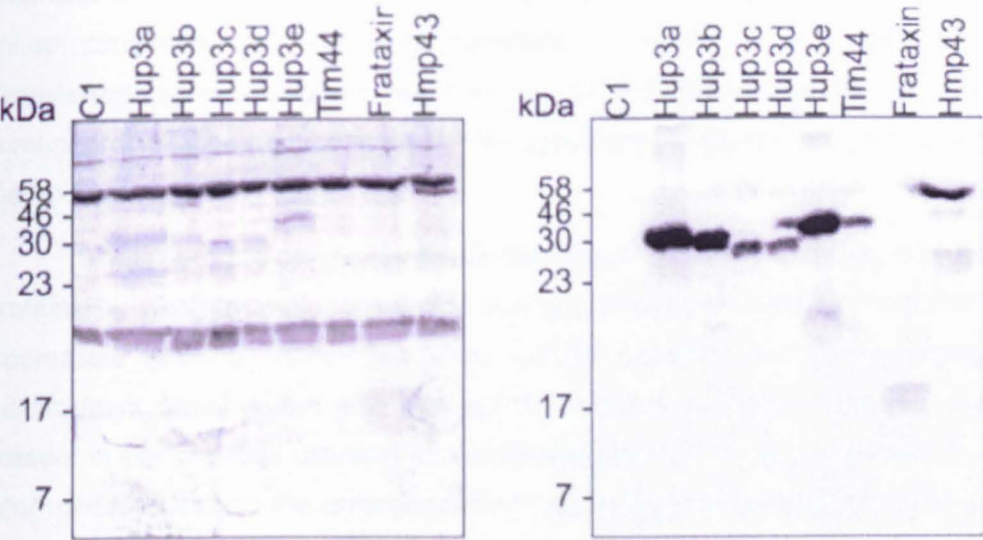


Figure 7.6.1 Co-immunoprecipitation of candidate proteins in 1% Triton X-100

Purified transformant hydrogenosomes were solubilised in 1% Triton X-100 solubilisation buffer, and any remaining organellar debris cleared by ultracentrifugation, the supernatant containing solubilised HA-recombinant protein was then recovered using an anti-HA antibody, protein-A Sepharose was then used to recover the immuno-complexed protein, and the Sepharose eluate was run on a 15% SDS PAGE gel. Duplicate gels were taken into western blotting (right) and protein detected by an anti-HA antibody.

The results shown in Figure 7.6.1 show recovery of the HA-tagged transformant protein, this is clearly visualised in both western detection and Coomassie staining. Immunological detection shows good sensitivity and reveals the presence of HA containing protein fragments (Hup3e, Hmp43) which are not visible on the Coomassie stained gel, or western blot control. In addition to the full length protein products, some higher molecular weight species are seen in Hup3a,e , and encouragingly show a similar size to complexes observed in BN PAGE and crosslinking experiments. This might suggest that these complexes exhibit robust stability which overcome the action of SDS.

The results of the co-immunoprecipitation are limited in the quantities of recovered protein. If native protein complexes are unstable over the time periods used to recover them from the hydrogenosomal lysate, the interacting proteins might be lost from the co-immunoprecipitation, and might result in an abundance of HA tagged protein but only residual quantities of interacting species, which might fall below the detection threshold for Coomassie detection. This effect might explain the limited recovery of additional molecular species seen in Figure 7.6.1.

To overcome the dynamic instability of transformant membrane complexes the crosslinking approach was redeveloped in tandem with co-immunoprecipitation. In this method, chemical crosslinking is used prior to hydrogenosome solubilisation to generate covalently linked complexes as previously observed in simple SDS PAGE/crosslinking analysis. Crosslinked hydrogenosomes are then solubilised and processed as before with the co-immunoprecipitation method to purify HA-tagged protein complexes. The results of this modified method are shown in Figure 7.6.2.

Similarly to the co-immunoprecipitation work duplicate samples were prepared to detect proteins by western immunodetection (using a primary antibody against the HA-tag), and by Coomassie staining. Immunodetection proved sensitive to detect unlinked monomers of transformant proteins, but also showed the addition of higher molecular weight species not present in the previous unlinked immunoprecipitation. This approach also has the benefit that information relating to the arrangement of interacting proteins with the HA-tagged is preserved within the crosslinked species.

These data would suggest that the crosslinked immunoprecipitation was recovering new molecular species, not recovered in unlinked co-immunoprecipitation, but similar to the unisolated complexes seen in BN PAGE and crosslinking experiments. These new intermolecular species represent immunopurified hydrogenosomal complexes which have been stabilised by chemical crosslinking. The results of the crosslink Co-IP do not however reproduce exactly the same pattern of species as been previously shown under crosslinking. Many species are missing, perhaps as a result of insolubility in Triton X-100 as opposed to SDS used to solubilise complexes in the crosslinking experiment, as such the number of intermolecular species in the crosslink Co-IP is fewer than in the original SDS crosslink experiment.

Complexes which are able to be resolved in the crosslink Co-IP which are not artefacts of cross reaction with the Anti-HA antibody (seen in the control) correspond to species observed in the crosslink experiments, particular the complexes between 60-80kDa which were also

observed in the BN PAGE and crosslink experiments, complexes at higher molecular weights are masked by secondary antibody cross reaction with the anti-HA antibody (seen at high molecular weight), but the intensity of this region is greater than the control, perhaps indicating the presence of additional crosslinking species in this mass region.

Coomassie detection is sensitive enough to detect the monomeric transformant proteins, but very poorly higher molecular weight species. Whilst visual detection of extra protein species is limited the eluates of the crosslink Co-IP are relatively enriched in these higher molecular weight species versus the original Co-IP experiments. Coomassie detection of hydrogenosomal proteins in the controls indicate that the Co-IP method highly purifies a small proportion of species from the original lysates, and thus eluates should only yield significant quantities of proteins which have been co-immunoprecipitated.

The protein species within the eluates represent the same species seen visually in the western blot, but without spatial resolution. As such eluates themselves can be analysed to determine protein constituents at the expense of mass resolution of complexes.

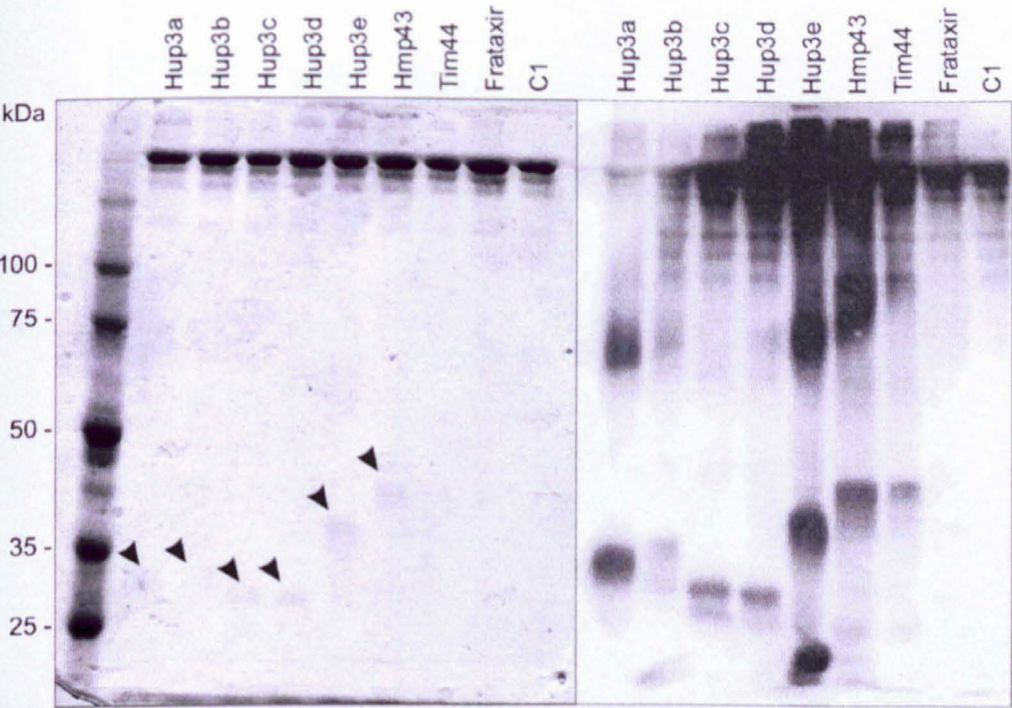


Figure 7.6.2 Crosslinking Co-IP of candidate proteins in 1% Triton X-100

Purified transformant hydrogenosomes were thawed and crosslinked as described previously (p75) and subsequently solubilised in a 1% Triton X-100 solubilisation buffer. Lysate was subsequently cleared by ultracentrifugation, and the supernatant transferred to fresh tubes. Recombinant protein was extracted by binding to an anti-HA antibody, before being pulled down with Protein-A Sepharose. 500 µg hydrogenosomal equivalent lysate was loaded into each lane above, one set of samples were stained with Coomassie (left) whilst the duplicated were detected by Western blotting. Recombinant protein was identified in the stained gel (arrows) with corresponding locations in the western

blot (right), in addition high molecular weight complexes are observed between 75-100kDa. These features are not present on Figure 7.6.1

7.7 Attempts to identify the protein interaction partners of the β -barrel translocase candidates

The techniques developed in the previous sections have indicated that the outer membrane β -barrel candidates form intermolecular species which have been demonstrated both within whole hydrogenosomal fractions, and also within purified co-immunoprecipitation eluates. The purification of unlinked and crosslinked complexes through immunoprecipitation provided an opportunity to directly identify interacting species through mass spectrometry analysis of electrophoretically resolved species. The following figures illustrate the attempts made to use mass spectrometry to identify interacting proteins.

7.7.1 Mass spectrometry combined with co-immunoprecipitation

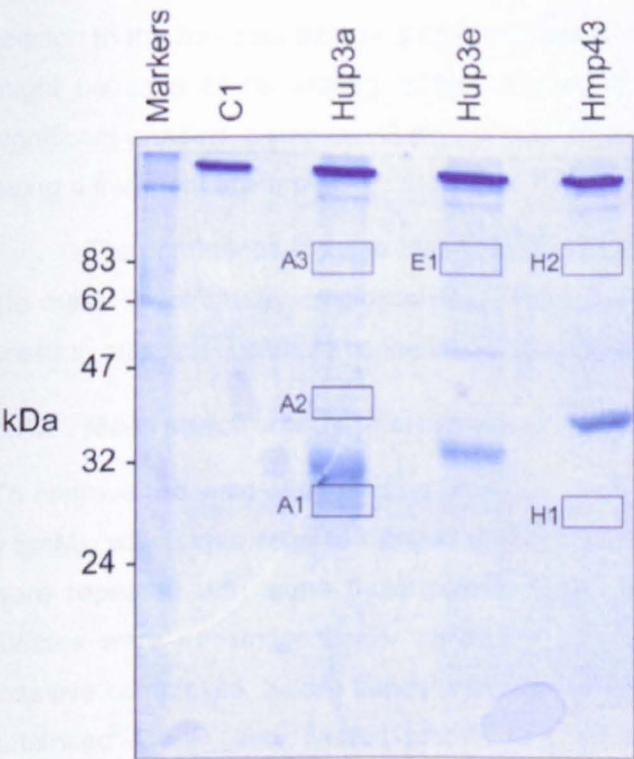


Figure 7.7.1 Extraction of co-immunoprecipitated species from selected β -barrel translocase candidates

Co-immunoprecipitation was used without crosslinking to extract native complexes from whole hydrogenosomal fractions. Proteins natively associating with the HA-tagged transformant protein were co-purified when these proteins were recovered using an anti-HA antibody. Co-IP eluates were then size separated under denaturing conditions in SDS PAGE, shown above. Portions of lanes occupied with additional species with respect to the parental strain C1 were excised (shown boxed) and submitted for mass spectrometry.

β -barrel proteins Hup3a,e and Hmp43 were selected for further study due to their ready solubilisation under co-immunoprecipitation conditions. purified hydrogenosomes from these transformants were solubilised in a 1% Triton X-100 solubilisation buffer without crosslinking, and transformant protein recovered using anti-HA antibody. Antibody and attached complexes were bound to Protein-A Sepharose beads, before elution with 0.1M glycine (pH2.8). Eluted proteins were size separated on a 12% SDS PAGE gel (Figure 7.7.1). The position of transformant protein is clearly seen on Coomassie staining, additional species were excised for mass spectrometry.

The excised fragments were then subject to trypsin digestion, and MALDI mass spectrometry, peptide fingerprints were then analysed in MASCOT. In general peptide recovery was low, with few peptides identified, and low coverage within the returned results. Samples A1, A3 returned the sequence for Hup3a, but no other significant results. Sample A2 returned results to adhesin protein AP51-3, and homologues. This protein is not known to have any relation to the hydrogenosome, and is thought to be expressed on the cell surface, its presence might be as a contaminating artifact. Sample E1, returned results for Hup3e but no other significant proteins. Samples H2 did not recover any meaningful spectra, but H1 was resolved to being a fragment of Hmp43.

The confidence in these results rests unsecurely on limited recovery of peptides, as such the mass spectrometry employed might have been insufficiently sensitive to detect all species present, alternately sample concentration might have been too low.

7.7.2 Mass spectrometry of crosslinked co-immunoprecipitation products

To improve the yield of interacting proteins, a crosslinking Co-IP was performed (using DSP at 0.5mM), which was seen to retrieve oligomeric complexes clearly seen in Figure 7.6.2. These were repeated with same three proteins Hup3a,e and Hmp43, due to their good recovery. Eluates were run under similar conditions on 10% SDS PAGE gels, to better resolve more massive complexes, before bands were excised for mass spectrometry. In contrast to previous 'un-linked' Co-IP, very limited protein was stained outside of the monomeric transformant protein. Excised fragments were digested and analysed as previously, however peptide recovery was extremely limited, and very poor confidence results were obtained, even from clearly Coomassie stained monomeric recombinant protein.

7.7.3 Analysis of mass spectrometry attempts

Immunodetection in previous Co-IP work positively identifies abundant quantities of both recombinant protein, crosslink Co-IPs are also shown to recover oligomeric complexes, as such these species are present and await analysis. The experience of mass spectrometry in this study has not however led to useful results.

Both unlinked and crosslinked reactions suffered from limited peptide recovery, this effect might be exacerbated in the crosslinked reactions due to lysine modification by DSP, which would inhibit trypsin digestion prior to MS analysis. Unfortunately even on DSP cleavage with DTT the residual modifications of lysine residues are likely to inhibit proteolysis. This problem could be circumvented by use of limited crosslinking (i.e. <0.5mM) or by use of alternate processing enzymes (for example Asp-N). More generally the peptide yield might be improved if eluates were submitted directly for MS analysis without SDS PAGE, this would prevent dispersal and dilution of protein, and maximise material available. However this approach would lose the 'spatial' information gained by SDS PAGE. Bands on the SDS PAGE correspond to unique combinations of proteins in a crosslinked complex, the complex composition could then be determined empirically by the mass and components present from the peptide fingerprint data from each band. Whilst this is sacrificed if eluates are directly sampled, a population of potentially interacting proteins would still be obtained.

Whilst the attempts of using MS were not fruitful in this study, with further investigation, especially with a direct examination of eluates would most likely quickly address which proteins interact with the candidate translocases, and suggest a population of proteins for *T. vaginalis* outer membrane translocases.

7.8 In summary

The molecular biology techniques employed in this section have revealed more about the real function of the translocase candidate than microscopy or bioinformatic analysis. These results indicate that all translocase candidates reside within the hydrogenosome- suggested by microscopy, and have definite and intimate relation with the hydrogenosomal membrane. These proteins have also been characterised with respect to their native complexes by multiple techniques, and have produced a variety of putative complexes, summarised in Table 7.1. These candidates identified from genomics data have complex functions in the hydrogenosomal membrane and are assembled into multimeric complexes. Analysis of the character of these complexes is continued next in the discussion. Finally the approaches shown in this section are

provide practical routes into the exploration of hydrogenosomal membrane complexes, whilst this project was unsuccessful in finishing this work with mass spectrometry, indications and alternate techniques suggest that the composition of outer membrane translocase complexes are very close to being revealed.

| Hup3a | Hup3b | Hup3c | Hup3d | Hup3e | Hmp43 | Tim44 | |
|-------|-------|-------|-------|-------|---------|-------|---------------|
| 47 | 55 | 57 | 47 | 47 | 60 | 45 | Cross-linking |
| 50 | | | | 60 | 85 | 60 | SDS PAGE |
| | | | | 65 | | 65 | |
| 75 | 75 | 75 | ≈65 | 75 | | 140 | |
| 140 | | | | 85 | | 150 | |
| 150 | | | | 150 | 140 | 175 | |
| | | | | | 240 (+) | | |
| 60 | 60 | 60 | 60 | 70 | 70 | | Cross-linking |
| | | | | | | | Co-IP |
| | 70* | | 70* | | 80* | | |
| 90 | 70 | 80 | 80 | 90 | 90 | | BN-PAGE |
| | 100 | | | | | | |
| | 120 | 120 | 140 | 120 | | | |

Table 7.1 Mass estimations for intermolecular species discovered in the biochemical analysis for *T. vaginalis* candidate translocases

This table summarises the intermolecular complexes determined by the different techniques employed in this chapter. Data from the SDS PAGE/crosslinking experiments is shown top in red, the results of crosslinking experiments followed by HA-tag immunoprecipitation are shown below (blue), and the complexes determined by BN PAGE are shown bottom (green). These data have been arranged to show commonalities between the proteins.

biochemical analysis of the translocase complexes in *T. vaginalis* has been carried out using a range of techniques including SDS PAGE, crosslinking, immunoprecipitation and BN PAGE. The results of these experiments are summarised in Table 7.1. The data show that the translocase complexes are composed of a number of proteins, with the most common being Hup3a, Hup3b, Hup3c, Hup3d, Hup3e, Hmp43 and Tim44. The data also show that the complexes are very stable, with many of the species being detected in both the crosslinking and immunoprecipitation experiments. This suggests that the complexes are not only stable but also have a high degree of specificity for the proteins involved. The data also show that the complexes are very similar in composition to those found in other species, such as *E. coli* and *S. typhimurium*, which further supports the idea that the translocase complexes are conserved across a wide range of species.

8 Discussion

In this section the critical points from both methodological approaches will be assembled together to highlight the key findings of this investigation. The findings from each chapter will be briefly recapitulated to detail the specific points which contribute to a final discussion on the nature of the preprotein import system in *T. vaginalis*. During this discussion, suggestions for further investigations to build upon this work will be introduced, especially with regard to the practical techniques developed here.

8.1 Analysis of the preprotein import assays

The import assay investigation was developed to pursue a “top down” perspective on the *T. vaginalis* membrane translocase machinery. It was hoped that this line of investigation would support previous work on the topic of preprotein interactions with isolated hydrogenosomes (Bradley 1997). To fulfil this, novel hydrogenosomal purification strategies, specifically using needle lysis as a simple and robust procedure to prepare lysates, and the use of iodixanol for tightly resolved organellar separation.

These approaches produced the necessary components of the *in vitro* import assay system, but failed to completely reproduce the previously observed kinetics with preproteins and isolated hydrogenosomes. However, some association of preprotein (adenylate kinase) with hydrogenosome fractions was discerned, and preprotein was protected from proteolysis in the presence of ATP, and eliminated when hydrogenosomal membranes were disrupted with Triton X-100 (Figure 4.4.2 p98).

Preprotein-hydrogenosome interactions were probed further using crosslinking as a technique to stabilise these associations. These species would be composed of the adenylate kinase probe, and proteins close enough to be crosslinked. Whilst some preproteins might be non-specifically crosslinked to random proteins, these should not form abundant populations. The observation of discrete species generated on crosslinking suggests that the preprotein was abundantly complexed with specific membrane targets. Whilst these might not necessarily represent a translocon, they nonetheless represent a protein with which the probe abundantly binds. Data from these investigations are briefly reproduced below in Figure 8.1.1.

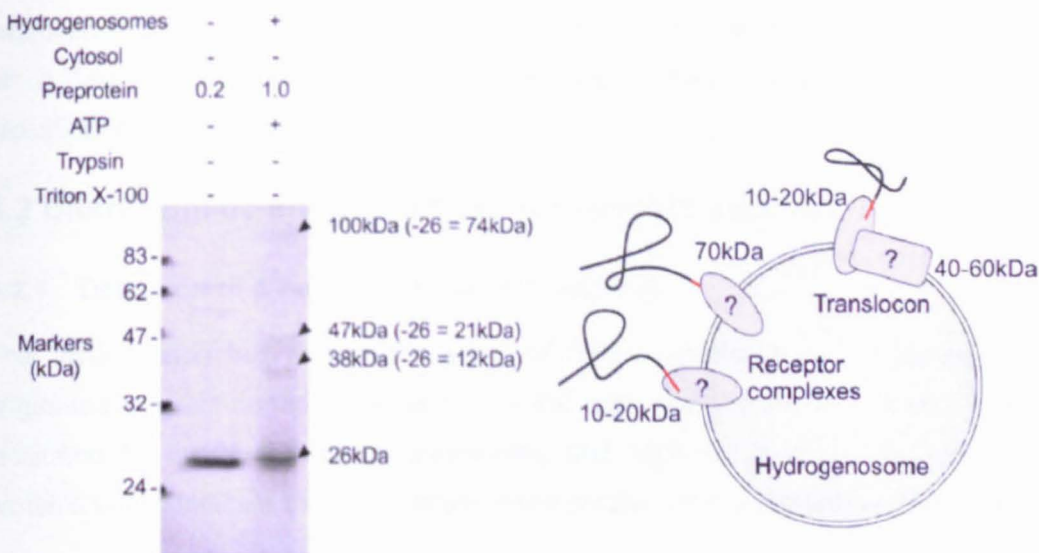


Figure 8.1.1 Identification of potential preprotein interaction complexes on the hydrogenosomal surface of *T. vaginalis*.

The data from the import assay chapter indicates that adenylate kinase preprotein associates with the surface of the translocon in discrete protein complexes. These could be visualized by chemically crosslinking probe and protein, and separating them electrophoretically. Three species were identified, of masses of approximately 12, 21, 74kDa, which are too small for canonical translocons (Mitochondrial TOM is >400kDa), but could represent receptor complexes.

Three putative species were detected by this crosslinking approach, and, assuming, that only one adenylate kinase molecule is within the complex, and that the species observed represent minimal combinations of proteins, the data suggest that the adenylate kinase probe might interact with species of ≈12, 21 and 74kDa.

Further analysis of these species can not be made from this data, but it is noteworthy that the principal receptors in mitochondria (Tom20, 22, 70) have similar masses to those predicted here. Other studies have shown that preproteins interact with isolated receptors immobilized upon resin columns (Brix et al. 1997), although preprotein: Tom40 interactions (predicted to have a mass of 65 kDa in *Trichomonas*, and thus which were not observed here) have been identified in previous mitochondrial work(Stan 2000).

It is clear that in this investigation the technique previously used to explore the kinetics of the preprotein import system could not identify translocons of similar size to the TOM complex (>400kDa), though smaller complexes were characterised. Attempts to purify these complexes using nickel bead affinity purification of the poly-His tag did not recover adequate yield for further analysis. However co-immunoprecipitation of the poly-His tag might prove superior in recovering these complexes, as demonstrated with Hup3 proteins, in addition different detergents appear amenable for the solubilisation of the *T. vaginalis* β-barrels and could be

explored for their efficacy in solubilising the translocon. These approaches could be conducted on a larger scale (>10mgs hydrogenosomes) without radio-labelled protein, to improve crosslinked species yield.

8.2 Bioinformatic analysis of the *T. vaginalis* genome

8.2.1 Detection of β -barrel proteins in *T. vaginalis*

This investigation has revealed a group of β -barrel proteins in *T. vaginalis*. These proteins originated in gram-negative bacteria, and the acquisition of these proteins in eukaryotes was facilitated by the events of endosymbiosis and organellogenesis. The presence of β -barrel proteins would indicate that *T. vaginalis* is the product of at least one endosymbiotic event.

Further analysis of the β -barrel proteins in *T. vaginalis* suggest some divergent relation to mitochondrial β -barrels, which would suggest a mitochondrial, or proto-mitochondrial ancestry for *T. vaginalis*. This investigation has identified a homologue to Sam50/Omp85, here named as Hmp43, and a group of proteins with similarity to Porins/Tom40 named Hup3a-e. The presence of these two families of β -barrels would suggest the potential for the *T. vaginalis* hydrogenosome to have the same kind of outer membrane architecture as the mitochondrion. The characterization of these β -barrels affirms the ubiquity of these proteins in eukaryotes, even in highly divergent eukaryotic clades. These results fit favourably with recent work discovering β -barrels, and β -translocases in other 'deep branching' eukaryotes (Maćasev et al. 2004; Dolezal et al. 2010).

This investigation briefly further characterized the *T. vaginalis* β -barrel proteins, identifying C-terminal β -barrel motifs which appear conserved in *T. vaginalis* and support this signal as a universal feature to eukaryotic β -barrels (Kutik et al. 2008). Domain and structural prediction of the *T. vaginalis* β -barrel proteins are not certain indicators of protein conformation, but do reinforce evidence that these proteins resemble other eukaryotic β -barrels, with respect to size, loop distribution, and strand number.

8.2.2 Detection of Tim17 domain containing proteins

Proteins containing the Tim17 domain have so far only been characterised in the inner membrane translocases of the mitochondrion. These proteins are thought to be of eukaryotic innovation, and are likely to have risen in response to a need to control and develop the inner membrane permeome of an endosymbiont (Gross & Bhattacharya 2009). This study identifies three proteins (Phats) in *T. vaginalis* with the Tim17 domain, and would represent the first

detection of the proteins within hydrogenosomes. These proteins were shown to contain the Tim17 domain and its characteristic hydrophobicity profile (especially with comparison to (Singha et al. 2008)). The radical divergence of the Phat proteins might be a local to the clades to which *T. vaginalis* belongs, as characterisation of a Tim17 homologue has been made in other Excavata, such as *T. brucei* (Singha et al. 2008), but not in closer related *G.intestinalis* (Jedelský et al. 2011).

Within this project a plasmid for Phatc was produced, but could not be successfully transformed, if this obstacle can be overcome, expression and microscopic analysis of this protein will help to complete a characterisation of the hydrogenosome. Investigation into the Phat family of proteins would complete a model for a translocase system in a hydrogenosome.

8.2.3 General points

Whilst it is not the purpose of this investigation to determine the deep branching nature of *T. vaginalis* it is noteworthy that phylogenetic analyses reproduce a similar picture for a variety of different proteins. Additionally when constructing phylogenetic analyses some relations to prokaryotic proteins were suggested.

In the case of Hmp43, which spans a region between Omp85/Sam50, and *T. vaginalis* Tim44, their particular relation between eukaryotic and prokaryotic groups might with further analysis, and more bikont sequence data, determine the point of divergence between *T. vaginalis* hydrogenosomes and mitochondria, as suggested .

Within the translocase system there seems to be remarkable differences in divergence of its proteins. This rate of divergence appears as a general outwards-in trend, whereby the inner most proteins, mtHsp70, PAM proteins, MPP, appear highly conserved, with divergence growing amongst the inner membrane proteins (Phats, Tim44), before almost unidentifiable homology in the outer membrane. Whether this trend is a general feature of divergent organelles would be an interesting question as are the driving factors to the different rates of divergence.

It might be speculated that certain properties of the outer membrane, especially the role of autocatalysis (Gross & Bhattacharya 2009) might make the outer membrane translocases more malleable to change, whilst even slight losses in efficiency to inner membrane machinery might lethally impair energy generation, and thus limiting divergence. In addition, the loss of transcriptional, translational activities from a degenerating genome-less mitochondrion might enable degeneration of outer membrane functions which previously had to support tRNA import.

8.3 Microscopy findings

8.3.1 Ultrastructural characterisation of *T. vaginalis*

Ultrastructural characterisation has been previously addressed in *T. vaginalis* by electron microscopy, which surpasses the inherent limits to resolution in optical microscopy; however within this project the optical resolution of compartments within *T. vaginalis*, a species rarely exceeding 15 microns does demonstrate that modern confocal microscopy can fulfil detailed studies into *T. vaginalis* ultrastructure.

Confocal microscopy has already been used to probe the distribution of particular markers, especially in the study of cytoskeleton, and Golgi (Benchimol et al. 2001) within *T. vaginalis*. In this study, off-the-shelf tracking agents have been successfully employed to visualise aspects of *T. vaginalis* cell biology. Whilst these studies represent a great illustration of the different types of localisation within the cell, it might also indicate that other molecular biology studies can approach *T. vaginalis* from a microscopic standpoint without exhaustive development of markers.

Bioinformatic strategies did identify a BIP homologue in *T. vaginalis*, and fixed cells did present reactivity with a commercial anti-BIP antibody. If these results genuinely correspond to a *T. vaginalis* BIP protein, then visualisation of the ER compartment can be visualised with a simple tracker. By similar token, ceramide based markers for the Golgi were able to visualise a compartment consistent with previous work using electron and confocal microscopy (Benchimol et al. 2001).

During the imaging study, the LysoTracker was found to localise particularly brightly within a limited number of small vesicles within the cell body. These vesicles are seen to be ≈ 1 micron or less, and seldom more than 5-6/cell, their brightness over background might represent marker abundance, or local pH (acidic). Difficulties in LysoTracker staining limited the quality of imaging, but the identity of these cryptic compartments might reveal a new compartment in *T. vaginalis*, possibly an acidic excretory one. Further characterization of the lysosome could combine the microscopy work initiated here together with proteomic analysis of the lysosomal fraction which we have been able to isolate by differential centrifugation (Figure 4.3.1 p95)

8.3.2 The *T. vaginalis* hydrogenosome

Several existing studies have been able to microscopically identify proteins and markers which specifically label the hydrogenosome. This study builds upon one of the previously studied

protein markers, but also develops the bioinformatically selected candidates as additional hydrogenosomal markers. In addition to existing matrix hydrogenosomal markers (frataxin) (Dolezal et al. 2007), this study has characterised marker proteins for the membrane compartments of the hydrogenosome, which show novel localisation.

The generation of membrane inserted marker proteins also facilitated direct measurement of the organelle from confocal images. Whilst the organelle density proved too crowded for image analysis programs to autonomously isolate and measure organelles, the marker proteins could be developed as a tool to examine the organelle population within *T. vaginalis*. The effects of drug treatment, specifically metronidazole has been shown influence the morphology of the hydrogenosome, the proteins characterised in this study could be developed to examine the effects of drug treatment on the hydrogenosomal population, using simple confocal microscopy, possibly in live cell studies. Approaching these questions from a confocal microscopy methodology might also prove more expedient than electron microscopy, especially with respect to quickly collecting data from many cells.

8.3.3 Hydrogenosomal localisation of hydrogenosomal candidates.

All successfully transformed candidate proteins were found to exhibit hydrogenosomal localization. Amongst the candidates two particular modes of localization were observed. Frataxin, the matrix marker was found to have a similar distribution to the homologue to mitochondrial Tim44. Whilst this latter protein is thought to reside on the inner membrane, the microscopy indicated some degree of luminal distribution, which was also confirmed through molecular biology approaches. However both frataxin and Tim44 proteins have a distribution that is distinct from the outer membrane proteins investigated (reviewed in Figure 6.5.2, p175).

Hmp43 and Hup3 proteins all similarly target the hydrogenosome, but exhibit different intra-organellar distribution. Together these candidate proteins were able to resolve the distribution not only of the hydrogenosome, within fixed cells, but also to discern different compartments within the organelle. These confocal images currently surpass other optical microscopy on hydrogenosomes.

Co-localisation studies using markers from different hydrogenosomal compartments could be used to assess protein distribution, and targeting through the organelle with suitably labelled preprotein probes. Similarly discrete translocon populations might be observed for proteins like the Phats through techniques such as FRET, and would reveal whether they operated like mitochondrial Tim17 domain translocases. Advances in super resolution

microscopy (for example Stimulated Emission Depletion microscopy (STED), or Structured Illumination Microscopy (SIM)) might also offer direct optical visualization of the composition of membrane translocases, the arrangements of translocons, and their membrane distribution.

8.4 Molecular biology characterisation

8.4.1 Candidate membrane localisation

Immunodetection of HA-tagged candidate proteins was proved to be consistent with microscopy findings, and confirmed that candidate localisation was to the hydrogenosome, and not lysosomal or other membranous compartment. The real value of the subcellular fractionation data was to build upon these findings and identify membrane localisation. The outer membrane β -barrel candidates were found to have a near exclusive localisation to the hydrogenosomal membranes. This is consistent with efficient insertion, especially as these proteins would be unstable without chaperones in the inter membrane space.

The inner membrane Tim44 homologue exhibited a less differentiated pattern of membrane localisation with some detectable sub-population present in the soluble fraction. These data might support the similarity in observations between this protein and frataxin made using confocal microscopy. In other organisms Tim44 has been characterised as a peripheral membrane protein which associates with the TIM23 translocon, whilst the majority of *T. vaginalis* Tim44 appears membrane associated, some remains soluble, perhaps an effect of preparation, or else over-expression, or disruption to association via the HA tag, alternately this might be a natural feature of this *T. vaginalis* protein, in any case the majority of protein is detected in membrane fractions consistent with this proteins function.

8.4.2 Putative membrane complexes

The molecular biology techniques developed for isolated hydrogenosomes were able to build upon localisation and begin to characterise membrane complexes. BN PAGE proved to be successful for outer membrane β -barrel candidates, and resolved distinct detergent specific complexes. Whilst BN PAGE relies on the careful extraction and electrophoretic separation of native complexes, the crosslinking technique developed for import assays was also able to approach these complexes through covalent crosslinking.

Hup3 family proteins were able to be resolved into three types of complex, small and intermediate complexes between 60kDa and 90kDa, and a larger complex around \approx 140-150kDa, and were detected both by BN PAGE and crosslinking/crosslinking Co-IP experiments.

These complexes are far smaller than typical TOM complexes (>400kDa, identified in (Model et al. 2001)) and must represent a different kind of organisation. Some TOM complex assembly intermediates are known with masses ≈ 100 kDa (Model et al. 2001; W Meisinger et al. 2007), and, given the presumed over expression of the putative β -barrel translocase relative to other factors this might lead to the accumulation of assembly intermediates. Alternately Hup3 family proteins might function as small oligomers, or else with few other proteins. Porins have been shown to act in this manner, and some work has characterised their oligomers in the Excavate *T.brucei* (Singha et al. 2009).

Hmp43 exhibits similar solubility to the Hup3 proteins in BN PAGE and exhibits a single ≈ 90 kDa complex, this complex is reproduced on crosslinking, though additional species are detected (Table 8.1). At the low mass end, ≈ 60 kDa, might again represent small oligomers, or assembly intermediates, however crosslinking does suggest larger ordered complexes which were not resolved on BN PAGE, these complexes at ≈ 140 , and ≈ 240 kDa might more adequately describe a SAM-like complex. These data make Hmp43 appear a promising candidate for the *T.vaginalis* SAM complex translocase.

| Reference complexes | | Hup3a | Hup3b | Hup3c | Hup3d | Hup3e | Hmp43 |
|--------------------------------|-----|-----------------------------|-------|-------|-------|-------|---------|
| | | 60 | 60 | 60 | 60 | 70 | |
| | | 47 | 55 | 57 | 47 | 47 | |
| | | 50 | | | | 60 | 60 |
| | | Small Hup3 complexes | | | | | |
| | | 70 | | | | | 70 |
| | | 75 | 75 | 75 | 75 | 85 | 80 |
| | | 70 | | | | | 85 |
| | | 90 | | 80 | 80 | 90 | 90 |
| | | 100 | | | | | |
| | | intermediate Hup3 complexes | | | | | Hmp43 |
| | | 120 | | | | | |
| | | 140 | | | | 150 | 140 |
| | | 150 | | | | | |
| | | large Hup3 complexes | | | | | Hmp43' |
| Assembly Intermediate of Tom40 | 100 | | | | | | |
| SAM core complex | 200 | | | | | | |
| Assembly intermediate of Tom40 | 250 | | | | | | 240 |
| SAM holocomplex | 350 | | | | | | Hmp43'' |
| TOM40 complex | 450 | | | | | | |

Table 8.1 Complexes determined for *T. vaginalis* β-barrel translocase candidates

Complexes determined by the different molecular biology approaches, BN PAGE (green), SDS PAGE + crosslink (red), SDS PAGE + crosslink + Co-IP (blue), have been arranged into groups by mass. These groups are shown against reference complexes characterized in literature for mitochondrial β-barrel proteins (W Meisinger et al. 2007). Hup3 proteins have been identified here as having three distinct groups, of low, intermediate and high molecular weight, small and large complexes are observed by two independent methods, and suggest that Hup3 proteins form multimeric complexes. Hmp43 is seen to have a distinct complex pattern with low molecular weight species, dominated by the mass of Hmp43, and two larger complexes Hmp43', Hmp43'' of much higher mass. The largest observed complex for Hmp43 compares favourably with the holocomplexes observed for the homologous Sam50.

The *T. vaginalis* homologue to Tim44 was unable to be resolved using BN PAGE, however crosslinking experiments were able to resolve a multitude of different intermolecular species, a result which would be consistent given Tim44's role as a bridge between two multimeric protein complexes. Whilst *T. vaginalis* proteins might be considerably different in length, prominent

crosslinked species match favourably with documented Tim44 complexes (Hutu et al. 2008) (Table 8.2). If the Tim44 homologue truly represents a reproduced part of the inner membrane translocase system, investigations into the Phat family of proteins might reveal common complexes, upon crosslinking.

| Documented Tim44 complexes | | <i>T. vaginalis</i> Tim44 crosslink species |
|----------------------------|-----|---|
| | | 45 |
| | | 60 |
| Tim44/Pam16/18 | 60 | 65 |
| TIM23 ^{CORE} | 90 | |
| Tim23' | 140 | 140 |
| | | 150 |
| TIM23Sort | 160 | |
| | | 175 |
| TIM23Sort' | 200 | |

Table 8.2 *T. vaginalis* Tim44 homologue crosslinked species

Whilst the *T. vaginalis* Tim44 homologue was not amenable to BN PAGE analysis, multiple crosslink species were generated, these species are presented here against previously characterized Tim44 complexes in *S. cerevisiae* (Hutu et al. 2008). The reference complexes illustrate the different roles of Tim44 in both J-complex (Tim44/Pam16/18) and Tim23 translocase (TIM23^{CORE}, Tim23'), but also its role as a bridge between the two (TIM23Sort, TIM23Sort')

8.4.3 Purification of hydrogenosomal protein complexes

BN PAGE and crosslinking provide a starting point for the analysis of these complexes and the molecular biology investigation was able to conclude with the establishment of protocols to purify these complexes from mixed hydrogenosomal fractions using co-immunoprecipitation. Whilst only a subset of the crosslinked species observed in the mixed hydrogenosomal fraction could be recovered by Co-IP, these complexes have been partially purified, and show that relatively few practical hurdles remain to isolating hydrogenosomal membrane complexes for mass spectroscopy analysis.

Initial attempts at mass spectrometry have yet to yield confident analyses on tested complexes, possibly resulting from residue modification by crosslinker which then might inhibit tryptic cleavage. However alternative mass spectroscopy strategies are possible for the future, including the use of LC-MS-MS (liquid chromatography tandem mass spectrometry) for analysis of the Co-IP elutions without subsequent SDS PAGE resolution. This approach might lose the information from the electrophoresis, especially with regards to the combination of proteins within a crosslinked species, but would overcome protein concentration issues and a population

of putative interaction partners. Whilst mass spectroscopy would have fulfilled a complete characterisation of the translocase candidates, the products of this investigation suggest encouraging roles for these proteins as components in multimeric membrane complexes.

8.5 In summary

This investigation aimed to complete gaps in our knowledge with respect to one aspect of the hydrogenosome's biology which has remained unanswered since the first work to characterise the import kinetics of this organelle.

This project pursued a whole systems approach to address preprotein translocase system. Genomic data not originally available to the researchers initially confronted with the hydrogenosome, was used to identify a series of candidate proteins. The results of this investigation suggest that a relatively complete preprotein translocase system can be composed from mitochondrially related proteins, and suggests that this is the architecture for preprotein import in the *T. vaginalis* hydrogenosome, an idea consistent with the ancestral derivation of these organelles from mitochondria (Embley et al. 2003).

This project has also developed practical approaches which were used to characterise translocase candidates, these techniques could be further employed to complete characterisation of the hydrogenosomal preprotein import system, with implications to define the evolutionary position of this organism. Furthermore the proteins characterised in this study in combination with advancing technology open a window on the *T. vaginalis* hydrogenosome with implications for the study of divergent organelles and also an essential system for a clinically relevant parasite.

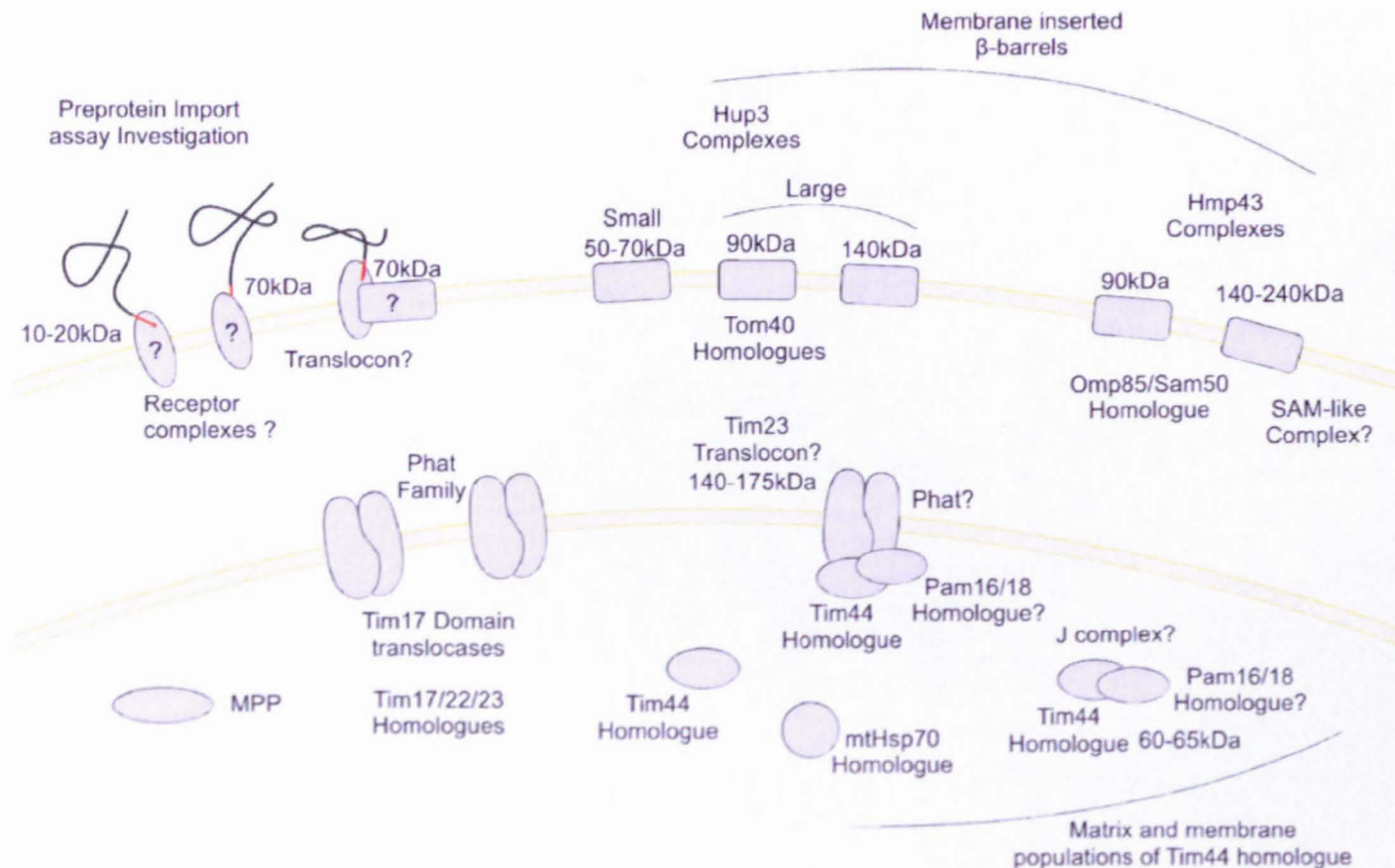


Figure 8.5.1 Proposed architecture for hydrogenosomal membrane translocase system based on the results of this investigation.

The data collected in this investigation has been presented in this figure to formulate a proposed membrane translocase system for the *T. vaginalis* hydrogenosome. This model does not definitively characterize the complexes detected in this study as representing complete translocons, but does show a proposed model for their distribution. This figure builds upon the complexes presented in Table 8.1 and Table 8.2

9 References

- Abramoff, M. D., Magalhaes, P. J., Ram, S. J. "Image processing with ImageJ". *Biophotonics International*, 11(7), pp. 36-42, 2004.
- Adams, K.L. & Palmer, J.D., 2003. Evolution of mitochondrial gene content: gene loss and transfer to the nucleus. *Molecular Phylogenetics and Evolution*, 29(3), pp.380-395.
- Aguilera, P., Barry, T. & Tovar, J., 2008. *Entamoeba histolytica* mitosomes: organelles in search of a function. *Experimental parasitology*, 118(1), pp.10–16.
- Ahting, U. et al., 2001. Tom40, the pore-forming component of the protein-conducting TOM channel in the outer membrane of mitochondria. *The Journal of Cell Biology*, 153(6), pp.1151-1160.
- Akhmanova, A. et al., 1998. A hydrogenosome with a genome. *Nature*, 396(6711), pp.527-528.
- Akhmanova, A. et al., 1999. A hydrogenosome with pyruvate formate-lyase: anaerobic chytrid fungi use an alternative route for pyruvate catabolism. *Molecular Microbiology*, 32(5), pp.1103-1114.
- Andersson, J.O. & Andersson, S.G., 1999. Insights into the evolutionary process of genome degradation. *Current Opinion in Genetics & Development*, 9(6), pp.664-671.
- Andersson, Siv G. E. et al., 2003. On the origin of mitochondria: a genomics perspective. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 358(1429), pp.165-179.
- Audia, J.P. & Winkler, H.H., 2006. Study of the five *Rickettsia prowazekii* proteins annotated as ATP/ADP translocases (Tlc): Only Tlc1 transports ATP/ADP, while Tlc4 and Tlc5 transport other ribonucleotides. *Journal of Bacteriology*, 188(17), pp.6261-6268.
- Bagos, P.G. et al., 2004. A Hidden Markov Model method, capable of predicting and discriminating beta-barrel outer membrane proteins. *BMC Bioinformatics*, 5:29.
- Bateman, A. et al., 2002. The Pfam protein families database. *Nucleic Acids Research*, 30(1), pp.276-280.
- Benchimol, M., 2009. Hydrogenosomes under microscopy. *Tissue and Cell*, 41(3), pp.151–168.

- Benchimol, M., 2008. The hydrogenosome peripheral vesicle: Similarities with the endoplasmic reticulum. *Tissue and Cell*, 40(1), pp.61–74.
- Benchimol, M., 2000. Ultrastructural characterization of the isolated hydrogenosome in *Tritrichomonas foetus*. *Tissue and Cell*, 32(6), pp.518–526.
- Benchimol, M., Johnson, P. J. & de Souza, W., 1996. Morphogenesis of the hydrogenosome: an ultrastructural study. *Biology of the Cell / Under the Auspices of the European Cell Biology Organization*, 87(3), pp.197-205.
- Benchimol, M. et al., 2001. Structure and division of the Golgi complex in *Trichomonas vaginalis* and *Tritrichomonas foetus*. *European Journal of Cell Biology*, 80(9), pp.593-607.
- Birnboim, H.C. & Doly, J., 1979. A rapid alkaline extraction procedure for screening recombinant plasmid DNA. *Nucleic Acids Research*, 7(6), pp.1513-1523.
- Boxma, B. et al., 2007. The [FeFe] hydrogenase of *Nyctotherus ovalis* has a chimeric origin. *BMC Evolutionary Biology*, 7(1), pp.230-242
- Bradley, P.J., 1997. Targeting and translocation of proteins into the hydrogenosome of the protist *Trichomonas*: similarities with mitochondrial protein import. *The EMBO Journal*, 16(12), pp.3484-3493.
- Bredemeier, R. et al., 2006. Functional and phylogenetic properties of the pore-forming beta-Barrel Transporters of the Omp85 Family. *Journal of Biological Chemistry*, 282(3), pp.1882-1890.
- Brix, J., Dietmeier, K. & Pfanner, N., 1997. Differential recognition of preproteins by the purified cytosolic domains of the mitochondrial import receptors Tom20, Tom22, and Tom70. *The Journal of Biological Chemistry*, 272(33), pp.20730-20735.
- Brown, D.M. et al., 1999. Alternative 2-keto acid oxidoreductase activities in *Trichomonas vaginalis*. *Molecular and Biochemical Parasitology*, 98(2), pp.203-214.
- Brown, M.T. et al., 2007. A functionally divergent hydrogenosomal peptidase with protomitochondrial ancestry. *Molecular Microbiology*, 64(5), pp.1154-1163.
- Brugerolle, G., 2004. Devescovinid features, a remarkable surface cytoskeleton, and epibiotic bacteria revisited in *Mixotricha paradoxa*, a parabasalid flagellate. *Protoplasma*, 224(1-2), pp.49-59.

- Bui, E.T. & Johnson, P J, 1996. Identification and characterization of [Fe]-hydrogenases in the hydrogenosome of *Trichomonas vaginalis*. *Molecular and Biochemical Parasitology*, 76(1-2), pp.305-310.
- Calvo, S.E. & Mootha, V.K., 2010. The mitochondrial proteome and human disease. *Annual Review of Genomics and Human Genetics*, 11(1), pp.25-44.
- Carlton, J.M. et al., 2007. Draft genome sequence of the sexually transmitted pathogen *Trichomonas vaginalis*. *Science*, 315(5809), pp.207-212.
- Cavalier-Smith, T., 2002. Chloroplast evolution: secondary symbiogenesis and multiple losses. *Current Biology*, 12(2), pp62–64.
- Cavalier-Smith, T., 1993. Kingdom protozoa and its 18 phyla. *Microbiology and Molecular Biology Reviews*, 57(4), p.953-994
- Cavalier-Smith, T., 2006. Origin of mitochondria by intracellular enslavement of a photosynthetic purple bacterium. *Proceedings of the Royal Society B: Biological Sciences*, 273(1596), pp.1943-1952.
- Cavalier-Smith, T., 2009. Predation and eukaryote cell origins: a coevolutionary perspective. *The International Journal of Biochemistry & Cell Biology*, 41(2), pp.307–322.
- Cavalier-Smith, T, 2002. The phagotrophic origin of eukaryotes and phylogenetic classification of Protozoa. *International Journal of Systematic and Evolutionary Microbiology*, 52(Pt 2), pp.297-354.
- Cerkasovová, A., Cerkasov, J. & Kulda, J., 1984. Metabolic differences between metronidazole resistant and susceptible strains of *Trichomonas foetus*. *Molecular and Biochemical Parasitology*, 11, pp.105-118.
- Chan, N. C. & Lithgow, T., 2008. The peripheral membrane subunits of the SAM complex function codependently in mitochondrial outer membrane biogenesis. *Molecular biology of the cell*, 19(1), pp.126-136.
- Chomicz, L. et al., 2004. Studies on the susceptibility of *Trichomonas hominis* to some abiotic factors. *Wiadomości Parazytologiczne*, 50(3), pp.405-409.
- Claros, M. G. & von Heijne, G., 1994. TopPred II: an improved software for membrane protein structure predictions. *Computer Applications in the Biosciences: CABIOS*, 10(6), pp.685-686.

- Cotch, M. F. et al., 1997. *Trichomonas vaginalis* associated with low birth weight and preterm delivery. *Sexually Transmitted Diseases*, 24(6), pp.353-360.
- Cuezva, J. M. et al., 1993. Molecular chaperones and the biogenesis of mitochondria and peroxisomes. *Biology of the Cell / Under the Auspices of the European Cell Biology Organization*, 77(1), pp.47-62.
- Dacks, J B & Doolittle, W.F., 2001. Reconstructing/deconstructing the earliest eukaryotes: how comparative genomics can help. *Cell*, 107(4), pp.419-425.
- Dacks, J B et al., 2002. Analyses of RNA Polymerase II genes from free-living protists: phylogeny, long branch attraction, and the eukaryotic big bang. *Molecular Biology and Evolution*, 19(6), pp.830-840.
- Danovaro, R. et al., 2010. The first metazoa living in permanently anoxic conditions. *BMC biology*, 8(1), pp.30-40
- Davidson, E. A. et al., 2002. An [Fe] hydrogenase from the anaerobic hydrogenosome-containing fungus *Neocallimastix frontalis* L2. *Gene*, 296(1-2), pp.45-52.
- Dembowski, M., 2001. Assembly of Tom6 and Tom7 into the TOM Core Complex of *Neurospora crassa*. *Journal of Biological Chemistry*, 276(21), pp.17679-17685.
- Diamond, L. S., 1957. The establishment of various trichomonads of animals and man in axenic cultures. *The Journal of Parasitology*, 43(4), pp.488-490.
- von Dohlen, C. D. et al., 2001. Mealybug beta-proteobacterial endosymbionts contain gamma-proteobacterial symbionts. *Nature*, 412(6845), pp.433-436.
- Dolezal, P. et al., 2010. The Essentials of protein import in the degenerate mitochondrion of *Entamoeba histolytica* D. Soldati-Favre, ed. *PLoS Pathogens*, 6(3), p.e1000812.
- Dolezal, P. et al., 2007. Frataxin, a conserved mitochondrial protein, in the hydrogenosome of *Trichomonas vaginalis*. *Eukaryotic Cell*, 6(8), pp.1431-1438.
- Doležal, P. et al., 2004. Malic enzymes of *Trichomonas vaginalis*: two enzyme families, two distinct origins. *Gene*, 329, pp.81-92.
- Drmotová, T. et al., 1996. Iron-ascorbate cleavable malic enzyme from hydrogenosomes of *Trichomonas vaginalis*: purification and characterization. *Molecular and Biochemical Parasitology*, 83(2), pp.221-234.

- Dubey, G. P. & Ben-Yehuda, S., 2011. Intercellular nanotubes mediate bacterial communication. *Cell*, 144(4), pp.590-600.
- Dyall, S. D. & Johnson P. J., 2004. Ancient invasions: from endosymbionts to organelles. *Science*, 304(5668), pp.253-257.
- Ellen, A. F. et al., 2009. Proteomic analysis of secreted membrane vesicles of archaeal *Sulfolobus* species reveals the presence of endosome sorting complex components. *Extremophiles: Life Under Extreme Conditions*, 13(1), pp.67-79.
- Embley T. M. & Martin, W. 2006. Eukaryotic evolution, changes and challenges. *Nature*, 440(7084), pp.623-630.
- Embley, T. M. et al., 2003. Mitochondria and hydrogenosomes are two forms of the same fundamental organelle. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 358(1429), pp.191-203.
- Embley, T. M. et al., 1995. Multiple origins of anaerobic ciliates with hydrogenosomes within the radiation of aerobic ciliates. *Proceedings. Biological Sciences / The Royal Society*, 262(1363), pp.87-93.
- Embley, T. M., 2006. Multiple secondary origins of the anaerobic lifestyle in eukaryotes. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 361(1470), pp.1055-1067.
- Emelyanov, V. V., 2007. Suggested mitochondrial ancestry of non-mitochondrial ATP/ADP. *Molekuliarnaia Biologiya*, 41(1), pp.59-70.
- Emelyanov, V. V., 2003. Mitochondrial connection to the origin of the eukaryotic cell. *European Journal of Biochemistry / FEBS*, 270(8), pp.1599-1618.
- Emelyanov, V. V. & Goldberg, A. V., 2011. Fermentation enzymes of *Giardia intestinalis*, pyruvate:ferredoxin oxidoreductase and hydrogenase, do not localize to its mitosomes. *Microbiology*, 157(Pt 6), pp.1602-1611.
- Esteban G, et al.1993. *Cyclidium porcatum* n.sp.: a free-living anaerobic scuticociliate containing a stable complex of hydrogenosomes, eubacteria and archaeobacteria. *European Journal of Protistology*, 29, pp.262–270.
- Felleisen, R. S., 1999. Host-parasite interaction in bovine infection with *Tritrichomonas foetus*. *Microbes and Infection / Institut Pasteur*, 1(10), pp.807-816.

- Forterre, P. & Philippe, H., 1999. Where is the root of the universal tree of life? *BioEssays*: 21(10), pp.871-879.
- Frazier, A. E. et al., 2004. Pam16 has an essential role in the mitochondrial protein import motor. *Nature Structural & Molecular Biology*, 11(3), pp.226-233.
- Funes, S. et al., 2011. Evolution of YidC/Oxa1/Alb3 insertases: three independent gene duplications followed by functional specialization in bacteria, mitochondria and chloroplasts. *Biological Chemistry*, 392(1-2), pp.13-19.
- Gabriel, K., Buchanan, S.K. & Lithgow, T., 2001. The alpha and the beta: protein translocation across mitochondrial and plastid outer membranes. *Trends in Biochemical Sciences*, 26(1), pp.36-40.
- Gabriel, K., Egan, B. & Lithgow, T., 2003. Tom40, the import channel of the mitochondrial outer membrane, plays an active role in sorting imported proteins. *The EMBO journal*, 22(10), pp.2380–2386.
- Garcia-Tamayo, J., Nuñez-Montiel, J. T. & de Garcia, H. P., 1978. An electron microscopic investigation on the pathogenesis of human vaginal trichomoniasis. *Acta Cytologica*, 22(6), pp.447-455.
- Germot, A., 1996. Presence of a mitochondrial-type 70-kDa heat shock protein in *Trichomonas vaginalis* suggests a very early mitochondrial endosymbiosis in eukaryotes. *Proceedings of the National Academy of Sciences*, 93(25), pp.14614-14617.
- Germot, A. & Philippe, H., 1999. Critical analysis of eukaryotic phylogeny: a case study based on the HSP70 family. *The Journal of Eukaryotic Microbiology*, 46(2), pp.116-124.
- Gevorkyan-Airapetov, L. et al., 2008. Interaction of Tim23 with Tim50 Is Essential for Protein Translocation by the Mitochondrial TIM23 Complex. *Journal of Biological Chemistry*, 284(8), pp.4865-4872.
- Glick, B. & Schatz, G., 1991. Import of Proteins into Mitochondria. *Annual Review of Genetics*, 25(1), pp.21-44.
- Gordon, D. M. et al., 1999. Maturation of frataxin within mammalian and yeast mitochondria: one-step processing by matrix processing peptidase. *Human Molecular Genetics*, 8(12), pp.2255-2262.

- de Graaf, R. M. et al., 2011. The organellar genome and metabolic potential of the hydrogen-producing mitochondrion of *Nyctotherus ovalis*. *Molecular Biology and Evolution*, 28 (8): pp2379-2391.
- Graham, J., 1994. The preparation of subcellular organelles from mouse liver in self-generated gradients of Iodixanol. *Analytical Biochemistry*, 220, pp.367-373.
- Gray, M. W., 1999. Mitochondrial evolution. *Science*, 283(5407), pp.1476-1481.
- Gray, M. W., 2005. Evolutionary biology: the hydrogenosome's murky past. *Nature*, 434(7029), pp.29-31.
- van Grinsven, K.W.A. et al., 2008. Acetate:succinate CoA-transferase in the hydrogenosomes of *Trichomonas vaginalis*: identification and characterization. *The Journal of Biological Chemistry*, 283(3), pp.1411-1418.
- Gromov, B. V., Seravin, L. N. & Gerasimova, Z. P., 1977. Bacteria-cortical symbionts of *Thichonympha turkestanica*, a protozoan from the gut tract of the termite *Hodotermes margabicus*. *Mikrobiologiya*, 46(5), pp.971-973.
- Gross, J. & Bhattacharya, D., 2009. Mitochondrial and plastid evolution in eukaryotes: an outsiders' perspective. *Nature Reviews Genetics*, 10(7), pp.495–505.
- Gross, J. & Bhattacharya, D., 2010. Uniting sex and eukaryote origins in an emerging oxygenic world. *Biology Direct*, 5:53.
- Habib, S. J., 2005. Assembly of the TOB complex of mitochondria. *Journal of Biological Chemistry*, 280(8), pp.6434-6440.
- Hall, T.A., 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT *Nucleic acids symposium series* 41, pp.95-98.
- Hampl, V. et al., 2008. Genetic evidence for a mitochondriate ancestry in the “amitochondriate” Flagellate *Trimastix pyriformis* R. Redfield, ed. *PLoS ONE*, 3(1), p.e1383.
- Heazlewood, J. L., 2004. Experimental analysis of the *Arabidopsis* mitochondrial proteome highlights signaling and regulatory components, provides assessment of targeting prediction programs, and indicates plant-specific mitochondrial proteins. *The Plant Cell Online*, 16(1), pp.241-256.
- Hedges, S. B. et al., 2004. A molecular timescale of eukaryote evolution and the rise of complex multicellular life. *BMC Evolutionary Biology*, 4:2.

- Hell, K. 2008. The Erv1-Mia40 disulfide relay system in the intermembrane space of mitochondria. *Biochimica Et Biophysica Acta*, 1783(4), pp.601-609.
- Herrmann, J. M., 2003. Converting bacteria to organelles: evolution of mitochondrial protein sorting. *Trends in Microbiology*, 11(2), pp.74-79.
- Hoffman P. F., et al., 1998. A neoproterozoic snowball earth. *Science*, 281(5381), pp.1342-1346.
- Holland H. D., 1994. Early proterozoic atmospheric change. *New York, N.Y: Columbia University Press*, pp.237-244.
- Horner, D. S., Hirt, R.P. & Embley, T. M., 1999. A single eubacterial origin of eukaryotic pyruvate: ferredoxin oxidoreductase genes: implications for the evolution of anaerobic eukaryotes. *Molecular Biology and Evolution*, 16(9), pp.1280-1291.
- Horner, D. S., Foster, P. G. & Embley, T. M., 2000. Iron hydrogenases and the evolution of anaerobic eukaryotes. *Molecular Biology and Evolution*, 17(11), pp.1695-1709.
- Hrdý, I & Müller, M., 1995. Primary structure and eubacterial relationships of the pyruvate:ferredoxin oxidoreductase of the amitochondriate eukaryote *Trichomonas vaginalis*. *Journal of Molecular Evolution*, 41(3), pp.388-396.
- Huson, D. H. et al., 2007. Dendroscope: An interactive viewer for large phylogenetic trees. *BMC Bioinformatics*, 8:460.
- Hutu, D. P. et al., 2008. Mitochondrial protein import motor: differential role of Tim44 in the recruitment of Pam17 and J-complex to the presequence translocase. *Molecular biology of the cell*, 19(6), pp.2642-2649
- Hwang, D. K. et al., 2007. Tim54p connects inner membrane assembly and proteolytic pathways in the mitochondrion. *The Journal of Cell Biology*, 178(7), pp.1161-1175.
- Ikeda-Ohtsubo, W. & Brune, A., 2009. Cospeciation of termite gut flagellates and their bacterial endosymbionts: *Trichonympha* species and "Candidatus Endomicrobium trichonymphae." *Molecular Ecology*, 18(2), pp.332-342.
- Jeanmougin, F. et al., 1998. Multiple sequence alignment with Clustal X. *Trends in Biochemical Sciences*, 23(10), pp.403-405.
- Jedelský, P. L. et al., 2011. The Minimal proteome in the reduced mitochondrion of the parasitic protist *Giardia intestinalis* B. Lightowlers, ed. *PLoS ONE*, 6(2), p.e17285.

- Jenkins, T. M. et al., 1991. Hydrogenosomal succinate thiokinase in *Tritrichomonas foetus* and *Trichomonas vaginalis*. *Biochemical and Biophysical Research Communications*, 179(2), pp.892-896.
- Jensen, R. E. & Johnson, A. E., 2001. Opening the door to mitochondrial protein import. *Nature Structural Biology*, 8(12), pp.1008-1010.
- Elfgrén I. K., 2003. Studies on the life cycles of akinete forming cyanobacteria, *Uppsala: Acta Universitatis Upsaliensis*.
- Keeling, P. J. & Palmer, J. D. 2008 Horizontal gene transfer in eukaryotic evolution *Nature Reviews Genetics* 9, pp605-618
- Kirschvink, J. L. et al., 2000. Paleoproterozoic snowball earth: extreme climatic and geochemical global change and its biological consequences. *Proceedings of the National Academy of Sciences of the United States of America*, 97(4), pp.1400-1405.
- Koebnik, R., Locher, K.P. & Van Gelder, P., 2000. Structure and function of bacterial outer membrane proteins: barrels in a nutshell. *Molecular Microbiology*, 37(2), pp.239-253.
- Koehler, C. M et al., 2000. Tim18p, a new subunit of the TIM22 complex that mediates insertion of imported proteins into the yeast mitochondrial inner membrane. *Molecular and cellular biology*, 20(4), pp.1187-1193.
- Kulda, J., Tachezy, J & Cerkasovová, A., 1993. *In vitro* induced anaerobic resistance to metronidazole in *Trichomonas vaginalis*. *The Journal of Eukaryotic Microbiology*, 40(3), pp.262-269.
- Kurland, C. G. & Andersson, S. G. E., 2000. Origin and evolution of the mitochondrial proteome. *Microbiology and molecular biology reviews*, 64(4), pp.786-820.
- Kutik, S. et al., 2008. Dissecting membrane insertion of mitochondrial beta-barrel proteins. *Cell*, 132(6), pp.1011-1024.
- Kvilekval, K. et al., 2009. Bisque: a platform for bioimage analysis and management. *Bioinformatics*, 26(4), pp.544-552.
- Kyte, J. & Doolittle, R.F., 1982. A simple method for displaying the hydropathic character of a protein. *Journal of Molecular Biology*, 157(1), pp.105-132.

- van der Laan, M. et al., 2005. Pam17 is required for architecture and translocation activity of the mitochondrial protein import motor. *Molecular and Cellular Biology*, 25(17), pp.7449-7458.
- Lantsman, Y. et al., 2008. Biochemical characterization of a mitochondrial-like organelle from *Blastocystis* sp. subtype 7. *Microbiology*, 154(9), pp.2757-2766.
- Lindmark, D. G. & Müller, M., 1973. Hydrogenosome, a cytoplasmic organelle of the anaerobic flagellate *Tritrichomonas foetus*, and its role in pyruvate metabolism. *The Journal of Biological Chemistry*, 248(22), pp.7724-7728.
- Lindsay, J. F., 2008. Was there a late Archean biospheric explosion? *Astrobiology*, 8(4), pp.823-839.
- Lloyd, D., Lindmark, D. G. & Müller, M., 1979. Adenosine triphosphatase activity of *Tritrichomonas foetus*. *Journal of General Microbiology*, 115(2), pp.301-307.
- Lloyd, D. & Kristensen, B., 1985. Metronidazole inhibition of hydrogen production in vivo in drug-sensitive and resistant strains of *Trichomonas vaginalis*. *Microbiology*, 131(4), pp.849-853.
- Lloyd, D., Ralphs, J.R. & Harris, J.C., 2002. *Giardia intestinalis*, a eukaryote without hydrogenosomes, produces hydrogen. *Microbiology*, 148(Pt 3), pp.727-733.
- Lockwood, B. C., North, M. J. & Coombs, G. H., 1988. The release of hydrolases from *Trichomonas vaginalis* and *Tritrichomonas foetus*. *Molecular and Biochemical Parasitology*, 30(2), pp.135-142.
- Mačasev, D. et al., 2004. Tom22', an 8-kDa trans-site receptor in plants and protozoans, is a conserved feature of the TOM complex that appeared early in the evolution of eukaryotes. *Molecular Biology and Evolution*, 21(8), pp.1557-1564.
- Mai, Z. et al., 1999. Hsp60 is targeted to a cryptic mitochondrion-derived organelle ("crypton") in the microaerophilic protozoan parasite *Entamoeba histolytica*. *Molecular and cellular biology*, 19(3), p.2198-2205.
- De Marais, D. J., 2000. Evolution. When did photosynthesis emerge on Earth? *Science*, 289(5485), pp.1703-1705.
- Marande, W. & Burger, G., 2007. Mitochondrial DNA as a genomic jigsaw puzzle. *Science*, 318(5849), p.415.

- Marande, W., Lukes, J. & Burger, G., 2005. Unique mitochondrial genome structure in diplomonids, the sister group of kinetoplastids. *Eukaryotic Cell*, 4(6), pp.1137-1146.
- Marchler-Bauer, A. et al., 2011. CDD: a Conserved Domain Database for the functional annotation of proteins. *Nucleic Acids Research*, 39(Database issue), pp.D225-229.
- Margulis L., 1970. Origin of eukaryotic cells. *Yale University Press*.
- Martin, W. & Müller, M., 1998. The hydrogen hypothesis for the first eukaryote. *Nature*, 392(6671), pp.37-41.
- Martin, W. & Koonin, E. V., 2006. Introns and the origin of nucleus–cytosol compartmentalization. *Nature*, 440(7080), pp.41-45.
- Martinez-Caballero, S. et al., 2006. Tim17p regulates the twin pore structure and voltage gating of the mitochondrial protein import complex TIM23. *Journal of Biological Chemistry*, 282(6), pp.3584-3593.
- Martínez-García, F. et al., 1996. Protozoan infections in the male genital tract. *The Journal of Urology*, 156(2 Pt 1), pp.340-349.
- McDougald, L. R., 2005. Blackhead disease (histomoniasis) in poultry: a critical review. *Avian Diseases*, 49(4), pp.462-476.
- Meisinger, C. et al., 2001. Protein import channel of the outer mitochondrial membrane: a highly stable Tom40-Tom22 core structure differentially interacts with preproteins, small tom proteins, and import receptors. *Molecular and Cellular Biology*, 21(7), pp.2337-2348
- Model, K. et al., 2001. Multistep assembly of the protein import channel of the mitochondrial outer membrane. *Nature Structural Biology*, 8(4), pp.361-370.
- Moreno, S. N. & Docampo, R., 1985. Mechanism of toxicity of nitro compounds used in the chemotherapy of trichomoniasis. *Environmental Health Perspectives*, 64, pp.199-208.
- Moro, F., 2001. Mitochondrial protein import: molecular basis of the ATP-dependent interaction of mtHsp70 with Tim44. *Journal of Biological Chemistry*, 277(9), pp.6874-6880.
- Muller, H. J., 1964. The relation of recombination to mutational advance *Research*, 106, pp.2-9.
- Müller, M., 1988. Energy metabolism of protozoa without mitochondria. *Annual Reviews in Microbiology*, 42(1), pp.465–488.
- Müller, M., 1993. Review Article: The hydrogenosome. *Microbiology*, 139(12), pp.2879-2889.

- Nielsen, M. H. & Nielsen, R., 1975. Electron microscopy of *Trichomonas vaginalis*: interaction with vaginal epithelium in human trichomoniasis. *Acta Pathologica Et Microbiologica Scandinavica. Section B, Microbiology*, 83(4), pp.305-320.
- Nixon, J. E. J. et al., 2003. Iron-dependent hydrogenases of *Entamoeba histolytica* and *Giardia lamblia*: activity of the recombinant entamoebic enzyme and evidence for lateral gene transfer. *The Biological Bulletin*, 204(1), pp.1-9.
- Noël, J.-C. et al., 2010. High prevalence of high-risk human papillomavirus infection among women with *Trichomonas vaginalis* infection on monolayer cytology. *Archives of Gynecology and Obstetrics*, 282(5), pp.503-505.
- Nowack, E. C. M. & Melkonian, M., 2010. Endosymbiotic associations within protists. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 365(1541), pp.699-712.
- O'Malley, M. A., 2010. The first eukaryote cell: an unfinished history of contestation. *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences*, 41(3), pp.212-224.
- Oppendoes, F. R. et al., 1984. Purification, morphometric analysis, and characterization of the glycosomes (microbodies) of the protozoan hemoflagellate *Trypanosoma brucei*. *The Journal of Cell Biology*, 98(4), pp.1178-1184.
- Özçelik s. et al., 2010 Investigation of the relationship between oral and dental health and presence of *Entamoeba gingivalis* and *Trichomonas tenax*. *Turkish Society for Parasitology*. 34(4) pp 155-159
- Paschen, S. A. et al., 2003. Evolutionary conservation of biogenesis of β -barrel membrane proteins. *Nature*, 426(6968), pp.862-866.
- Paul, R. G., Williams, A. G. & Butler, R. D., 1990. Hydrogenosomes in the rumen entodiniomorphid ciliate *Polyplastron multivesiculatum*. *Microbiology*, 136(10), pp.1981-1989.
- Peixoto, P. M. V. et al., 2007. Awakening TIM22, a dynamic ligand-gated channel for protein insertion in the mitochondrial inner membrane. *Journal of Biological Chemistry*, 282(26), pp.18694-18701.
- Petrakis, N., Alcock, F. & Tokatlidis, K., 2009. Mitochondrial ATP-independent chaperones. *IUBMB Life*, 61(9), pp.909-914.

- Plümper, E., Bradley, P.J. & Johnson, P J, 2000. Competition and protease sensitivity assays provide evidence for the existence of a hydrogenosomal protein import machinery in *Trichomonas vaginalis*. *Molecular and Biochemical Parasitology*, 106(1), pp.11-20.
- Van der Pol, B., 2007. *Trichomonas vaginalis* infection: the most prevalent nonviral sexually transmitted infection receives the least public health attention. *Clinical Infectious Diseases: An Official Publication of the Infectious Diseases Society of America*, 44(1), pp.23-25.
- Poole, A.M., 2006. Did group II intron proliferation in an endosymbiont-bearing archaeon create eukaryotes? *Biology Direct*, 1:36.
- Popot, J. L. & de Vitry, C., 1990. On the microassembly of integral membrane proteins. *Annual Review of Biophysics and Biophysical Chemistry*, 19, pp.369-403.
- Quon, D. V., d' Oliveira, C. E. & Johnson, P. J., 1992. Reduced transcription of the ferredoxin gene in metronidazole-resistant *Trichomonas vaginalis*. *Proceedings of the National Academy of Sciences of the United States of America*, 89(10), p.4402.
- Race, H. L., Hermann, R. G. & Martin, W., 1999. Why have organelles retained genomes? *Trends in Genetics* 15(9), pp.364-370.
- Reinders, J. et al., 2006. Toward the complete yeast mitochondrial proteome: Multidimensional separation techniques for mitochondrial proteomics. *Journal of Proteome Research*, 5(7), pp.1543-1554.
- Remmert, M. et al., 2010. Evolution of outer membrane beta-barrels from an ancestral beta beta hairpin. *Molecular Biology and Evolution*, 27(6), pp.1348-1358.
- Retief, J.D., 2000. Phylogenetic analysis using PHYLIP. *Methods in Molecular Biology (Clifton, N.J.)*, 132, pp.243-258.
- Rew, R. S. & Saz, H. J., 1974. Enzyme localization in the anaerobic mitochondria of *Ascaris lumbricoides*. *The Journal of Cell Biology*, 63(1), pp.125-135.
- Rodríguez, M. A. et al., 1998. The pyruvate:ferredoxin oxidoreductase enzyme is located in the plasma membrane and in a cytoplasmic structure in *Entamoeba*. *Microbial Pathogenesis*, 25(1), pp.1-10.
- Rosenthal, B. et al., 1997. Evidence for the bacterial origin of genes encoding fermentation enzymes of the amitochondriate protozoan parasite *Entamoeba histolytica*. *Journal of Bacteriology*, 179(11), pp.3736-3745.

- Sakharkar, K. R., 2004. Genome reduction in prokaryotic obligatory intracellular parasites of humans: a comparative analysis. *International Journal of Systematic and Evolutionary Microbiology*, 54(6), pp.1937-1941.
- Sakuragi, S., Liu, Q. & Craig, E., 1999. Interaction between the nucleotide exchange factor Mge1 and the mitochondrial Hsp70 Ssc1. *The Journal of Biological Chemistry*, 274(16), pp.11275-11282.
- Sassera, D. et al., 2006. "*Candidatus Midichloria mitochondrii*", an endosymbiont of the tick *Ixodes ricinus* with a unique intramitochondrial lifestyle. *International Journal of Systematic and Evolutionary Microbiology*, 56(11), pp.2535-2540.
- Schiller, D. et al., 2008. Residues of Tim44 involved in both association with the translocon of the inner mitochondrial membrane and regulation of mitochondrial Hsp70 tethering. *Molecular and Cellular Biology*, 28(13), pp.4424-4433.
- Schiller, D., 2009. Pam17 and Tim44 act sequentially in protein import into the mitochondrial matrix. *The International Journal of Biochemistry & Cell Biology*, 41(11), pp.2343-2349.
- Schleiff, E. & Becker, T., 2011. Common ground for protein translocation: access control for mitochondria and chloroplasts. *Nature Reviews. Molecular Cell Biology*, 12(1), pp.48-59.
- Schleiff, E. & Soll, J., 2005. Membrane protein insertion: mixing eukaryotic and prokaryotic concepts. *EMBO reports*, 6(11), pp.1023-1027.
- Schmitt, S., 2005. Role of Tom5 in maintaining the structural stability of the TOM complex of mitochondria. *Journal of Biological Chemistry*, 280(15), pp.14499-14506.
- Schwarz, E. & Neupert, W., 1994. Mitochondrial protein import: mechanisms, components and energetics. *Biochimica Et Biophysica Acta*, 1187(2), pp.270-274.
- Searcy, D. G., 2003. Metabolic integration during the evolutionary origin of mitochondria. *Cell research*, 13(4), pp.229-238.
- Singha, U. K. et al., 2008. Characterization of the mitochondrial inner membrane protein translocator Tim17 from *Trypanosoma brucei*. *Molecular and biochemical parasitology*, 159(1), pp.30-43.
- Sirrenberg, C. et al., 1996. Import of carrier proteins into the mitochondrial inner membrane mediated by Tim22. *Nature*, 384(6609), pp.582-585.

- Sogin, M. L., 1991. Early evolution and the origin of eukaryotes. *Current Opinion in Genetics & Development*, 1(4), pp.457-463.
- Sorvillo, F. et al., 2001. *Trichomonas vaginalis*, HIV, and African-Americans. *Emerging Infectious Diseases*, 7(6), pp.927-932.
- Stan, T., 2000. Recognition of preproteins by the isolated TOM complex of mitochondria. *The EMBO Journal*, 19(18), pp.4895-4902.
- Stechmann, A., et al. 2008. Organelles in *Blastocystis* that blur the distinction between mitochondria and hydrogenosomes. *Current Biology: CB*, 18(8), pp.580-585.
- Stefanic, S. et al., 2006. Organelle proteomics reveals cargo maturation mechanisms associated with Golgi-like encystation vesicles in the early-diverged protozoan *Giardia lamblia*. *The Journal of Biological Chemistry*, 281(11), pp.7595-7604.
- Steinbüchel, A. & Müller, M., 1986. Anaerobic pyruvate metabolism of *Tritrichomonas foetus* and *Trichomonas vaginalis* hydrogenosomes. *Molecular and Biochemical Parasitology*, 20(1), pp.57-65.
- Summons, R. E. et al., 1999. 2-Methylhopanoids as biomarkers for cyanobacterial oxygenic photosynthesis. *Nature*, 400(6744), pp.554-557.
- Suzuki, H., Maeda, M. & Mihara, K., 2002. Characterization of rat TOM70 as a receptor of the preprotein translocase of the mitochondrial outer membrane. *Journal of cell science*, 115(9), pp.1895-1905.
- Syvanen, M., 1985. Cross-species gene transfer; implications for a new theory of evolution. *Journal of Theoretical Biology*, 112(2), pp.333-343.
- Tachezy, J, Sánchez, L. B. & Müller, M., 2001. Mitochondrial type iron-sulfur cluster assembly in the amitochondriate eukaryotes *Trichomonas vaginalis* and *Giardia intestinalis*, as indicated by the phylogeny of IscS. *Molecular Biology and Evolution*, 18(10), pp.1919-1928.
- Takamiya, S. et al., 1994. Respiratory chain of the lung fluke *Paragonimus westermani*: facultative anaerobic mitochondria. *Archives of Biochemistry and Biophysics*, 312(1), pp.142-150.
- Tang, L. F., Watanabe, I. & Liu, C. C., 1990. Limited multiplication of symbiotic cyanobacteria of *Azolla* spp. on artificial media. *Applied and Environmental Microbiology*, 56(11), pp.3623-3626.

- Terada, K. et al., 1996. The requirement of heat shock cognate 70 protein for mitochondrial import varies among precursor proteins and depends on precursor length. *Molecular and cellular biology*, 16(11), pp.6103-6109.
- Thrash, J.C. et al., 2011. Phylogenomic evidence for a common ancestor of mitochondria and the SAR11 clade. *Scientific Reports* 1:13
- Tjaden, J. et al., 2004. A divergent ADP/ATP carrier in the hydrogenosomes of *Trichomonas gallinae* argues for an independent origin of these organelles. *Molecular Microbiology*, 51(5), pp.1439-1446.
- Tovar, J., Fischer, A. & Clark, C G, 1999. The mitosome, a novel organelle related to mitochondria in the amitochondrial parasite *Entamoeba histolytica*. *Molecular Microbiology*, 32(5), pp.1013-1021.
- Truscott, K. N et al., 2003. A J-protein is an essential subunit of the presequence translocase-associated protein import motor of mitochondria. *The Journal of cell biology*, 163(4), pp.707-713.
- Turner, G. & Müller, M., 1983. Failure to detect extranuclear DNA in *Trichomonas vaginalis* and *Tritrichomonas foetus*. *The Journal of Parasitology*, 69(1), pp.234-236.
- Vaidya, A. B. & Mather, M. W., 2009. Mitochondrial evolution and functions in malaria parasites. *Annual Review of Microbiology*, 63, pp.249-267.
- Voncken, F. G. et al., 2002. A hydrogenosomal [Fe]-hydrogenase from the anaerobic chytrid *Neocallimastix* sp. L2. *Gene*, 284(1-2), pp.103-112.
- Wanner, G., Vogl, K. & Overmann, J., 2008. Ultrastructural characterization of the prokaryotic symbiosis in "*Chlorochromatium aggregatum*." *Journal of Bacteriology*, 190(10), pp.3721-3730.
- Wideman, J.G. et al., 2010. Roles of the Mdm10, Tom7, Mdm12, and Mmm1 proteins in the assembly of mitochondrial outer membrane proteins in *Neurospora crassa*. *Molecular Biology of the Cell*, 21(10), p.1725.
- Windsor, J. J. & Johnson, E. H., 1999. *Dientamoeba fragilis*: the unflagellated human flagellate. *British Journal of Biomedical Science*, 56(4), pp.293-306.
- Wolff, E. D. S. et al., 2009. Common avian infection plagued the tyrant dinosaurs. *PloS One*, 4(9), p.e7288.

- Wujek D. E., 1979. Intracellular bacteria in the blue-green Alga *Pleurocapsa minor*. *Transactions of the American Microscopical Society*, 98(1), pp.143-145.
- Yamamoto, H. et al., 2009. Roles of Tom70 in import of presequence-containing mitochondrial proteins. *The Journal of Biological Chemistry*, 284(46), pp.31635-31646.
- Yamano, K., 2005. The phosphate carrier has an ability to be sorted to either the TIM22 pathway or the TIM23 pathway for its import into yeast mitochondria. *Journal of Biological Chemistry*, 280(11), pp.10011-10017.
- Young, J. C., Hoogenraad, N. J. & Hartl, F. U., 2003. Molecular chaperones Hsp90 and Hsp70 deliver preproteins to the mitochondrial import receptor Tom70. *Cell*, 112(1), pp.41-50.
- Yubuki, N. et al., 2009. Ultrastructure and molecular phylogeny of *Calkinsia aureus*: cellular identity of a novel clade of deep-sea euglenozoans with epibiotic bacteria. *BMC microbiology*, 9(1), p.16-38.
- Yutin, N. et al., 2009. The origins of phagocytosis and eukaryogenesis. *Biology Direct*, 4:9.

10 Appendices and supplementary data

| | | |
|---------|--|-----|
| 10.1.1 | Appendix 1 Polyphyly of mitochondrially derived organelles..... | 235 |
| 10.2.1 | Appendix 2 Suppliers of reagents used in this study | 239 |
| 10.2.2 | Appendix 3 pTagVag2 vector, and generation of Tv. Tim44 containing construct..... | 241 |
| 10.2.3 | Appendix 4 Generation of the β -barrel containing expression vectors from pTagVag2 | 242 |
| 10.2.4 | Appendix 5 Generation of constructs expressing Tv. ferredoxin and adenylate kinase | 243 |
| 10.3.1 | Appendix 6 Tom70 like proteins in <i>T. vaginalis</i> | 244 |
| 10.3.2 | Appendix 7 transmembrane prediction of Hup3 proteins | 245 |
| 10.3.3 | Appendix 8 transmembrane prediction of Hmp43..... | 246 |
| 10.3.4 | Appendix 9 PAM homologues in <i>T.vaginalis</i> | 247 |
| 10.3.5 | Appendix 10 TopPred topology prediction of <i>S. cerevisiae</i> Tim17 superfamily proteins | 248 |
| 10.3.6 | Appendix 11 Detection of multiple Hsp70 homologues..... | 249 |
| 10.3.7 | Appendix 12 Hsp70-like protein families in <i>T. vaginalis</i> | 250 |
| 10.3.8 | Appendix 13 Identification of mtHsp70 homologues..... | 251 |
| 10.4.1 | Appendix 14 Confocal analysis of DAPI stained cells..... | 252 |
| 10.4.2 | Appendix 15 Confocal microscopy of PI stained <i>T. vaginalis</i> cells | 253 |
| 10.4.3 | Appendix 16 A 3D reconstruction of a Golgi apparatus | 253 |
| 10.4.4 | Appendix 17 Epifluorescent images of Transformant Hup3d | 254 |
| 10.4.5 | Appendix 18 3D reconstruction of frataxin transformant cells | 256 |
| 10.4.6 | Appendix 19 3D reconstruction of Tim44 transformant cells..... | 257 |
| 10.4.7 | Appendix 20 Distribution of HA- Hup3a..... | 258 |
| 10.4.8 | Appendix 21 Intracellular distribution of Hup3c | 259 |
| 10.4.9 | Appendix 22 Intracellular distribution of Hup3d | 260 |
| 10.4.10 | Appendix 23 Intracellular distribution of LysoTracker bright bodies..... | 261 |
| 10.4.11 | Appendix 24 measuring the dimensions of the hydrogenosome | 262 |

10.1 Appendices for Chapter 1: Introduction

10.1.1 Appendix 1 Polyphyly of mitochondrially derived organelles

This appendix presents some of the variety in mitochondrially derived organelles. Literature sources were collected and presented in the following table, and taxonomic cladogram (produced using the ITOL taxonomic database). A bibliography of references used to compile the figures for this appendix has also been added after the diagram.

| Organism | Organelle | Reference |
|-----------------------------------|--|--------------------------|
| <i>Blastocystis</i> | mitochondrial-like organelle | (Stechmann et al. 2008) |
| <i>Breviata anathema</i> | mitochondrial relict proteins | (Minge et al. 2009) |
| <i>Carpediemonas</i> | hydrogenosome-like | (Simpson et al. 2002) |
| <i>Cryptosporidium</i> | mitsome like | (Riordan et al. 2003) |
| <i>Cyclidium glaucoma</i> | mitochondria | (Esteban et al 1993) |
| <i>Cyclidium porcatum</i> | hydrogenosome | (Esteban et al 1993) |
| <i>Dasytricha</i> | hydrogenosome | (Yarlett et al. 1982) |
| <i>ruminantium</i> | | |
| <i>Encephalitozoon cuniculi</i> | mitosome | (Tsaousis et al. 2008) |
| <i>Entamoeba</i> | mitosomes | (Tovar et al. 1999) |
| <i>Giardia</i> | mitosomes | (Tovar et al. 2003) |
| <i>Jakoba</i> | mitochondrial relict proteins | (Simpson et al. 2006) |
| <i>Mastigamoeba</i> | mitochondria-like | (Gill et al. 2007) |
| <i>Naegleria fowleri</i> | mitochondria/mitochondria-like | (Horner et al. 1996) |
| <i>Neocallimastix</i> | hydrogenosome | (Davidson et al. 2002) |
| <i>Nyctotherus ovalis</i> | mitochondria-like/hydrogenosome | (de Graaf et al. 2011) |
| <i>Pelomyxa</i> | none characterised, theorised secondary loss | (Edgcomb et al. 2002) |
| <i>Perkinsus</i> | mitochondria | (Sunila et al. 2001) |
| <i>Piromyces</i> | hydrogenosome | (Akhmanova et al. 1999) |
| <i>Plagiopyla nasuta</i> | relict organelles | (Fenchel et al. 1977) |
| <i>Psalteriomonas lanterna</i> | hydrogenosome | (de Graaf et al. 2009) |
| <i>Trachipleistophora hominis</i> | mitosome | (Goldberg et al. 2008) |
| <i>Trichomonas</i> | hydrogenosome | (Lindmark & Müller 1973) |
| <i>Trimastix pyriformis</i> | anaerobic mitochondria | (Hampl et al. 2008) |
| <i>Trimyema compressum</i> | hydrogenosome | (Finlay et al. 1993) |

Table 10.1 Organisms harboring mitochondrially derived organelles.

The variety of species containing mitochondrial derived organelles is demonstrated in the table above. In each case each organism is presented with a description of organelle type, and reference.

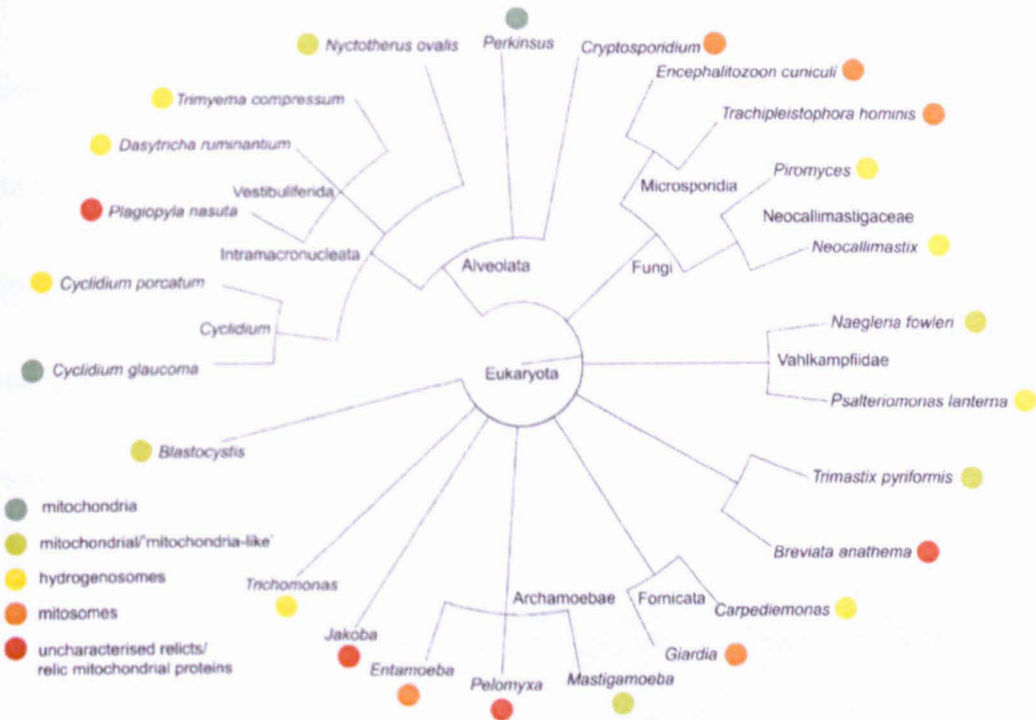


Figure 110.1.1.2 taxonomic representation of mitochondrially derived organelles.

The ITOL taxonomic server (<http://itol.embl.de/>) was used to construct the above figure showing the polyphyly of species containing mitochondrially derived organelles. This figure shows that mitochondrially derived organelles are found in a diverse number of taxa, as well as occurring within otherwise mitochondriate genus (for example *Cyclidium*).

References for Appendix 1.

Akhmanova, A. et al., 1999. A hydrogenosome with pyruvate formate-lyase: anaerobic chytrid fungi use an alternative route for pyruvate catabolism. *Molecular Microbiology*, 32(5), pp.1103-1114.

Davidson, E.A. et al., 2002. An [Fe] hydrogenase from the anaerobic hydrogenosome-containing fungus *Neocallimastix frontalis* L2. *Gene*, 296(1-2), pp.45-52.

Edgcomb, V.P. et al., 2002. Pelobionts are degenerate protists: insights from molecules and morphology. *Molecular Biology and Evolution*, 19(6), pp.978-982.

Esteban G, Guhl B.E, Ken J.C, Embley T.M, Finlay B.J., 1993. *Cyclidium porcatum* n.sp.: a free-living anaerobic scuticociliate containing a stable complex of hydrogenosomes, eubacteria and archaeobacteria. *Eur. J. Protistol.*, 29, pp.262-270.

Fenchel, T., Perry, T. & Thane, A., 1977. Anaerobiosis and symbiosis with bacteria in free-living ciliates. *The Journal of Protozoology*, 24(1), pp.154-163.

- Finlay, B.J., Embley, T M & Fenchel, T., 1993. A new polymorphic methanogen, closely related to *Methanocorpusculum parvum*, living in stable symbiosis within the anaerobic ciliate *Trimyema* sp. *Journal of General Microbiology*, 139(2), pp.371-378.
- Gill, E.E. et al., 2007. Novel mitochondrion-related organelles in the anaerobic amoeba *Mastigamoeba balamuthi*. *Molecular Microbiology*, 66(6), pp.1306-1320.
- Goldberg, A.V. et al., 2008. Localization and functionality of microsporidian iron-sulphur cluster assembly proteins. *Nature*, 452(7187), pp.624-628.
- de Graaf, R. M. et al., 2011. The organellar genome and metabolic potential of the hydrogen-producing mitochondrion of *Nyctotherus ovalis*. *Molecular Biology and Evolution*.
- de Graaf, Rob M et al., 2009. The hydrogenosomes of *Psalteriomonas lanterna*. *BMC Evolutionary Biology*, 9, p.287.
- Hampl, V. et al., 2008. Genetic evidence for a mitochondriate ancestry in the "amitochondriate" Flagellate *Trimastix pyriformis* R. Redfield, ed. *PLoS ONE*, 3(1), p.e1383.
- Horner, D S et al., 1996. Molecular data suggest an early acquisition of the mitochondrion endosymbiont. *Proceedings. Biological Sciences / The Royal Society*, 263(1373), pp.1053-1059.
- Lindmark, D.G. & Müller, M, 1973. Hydrogenosome, a cytoplasmic organelle of the anaerobic flagellate *Tritrichomonas foetus*, and its role in pyruvate metabolism. *The Journal of Biological Chemistry*, 248(22), pp.7724-7728.
- Minge, M.A. et al., 2009. Evolutionary position of breviate amoebae and the primary eukaryote divergence. *Proceedings of the Royal Society B: Biological Sciences*, 276(1657), pp.597-604.
- Riordan, C.E. et al., 2003. *Cryptosporidium parvum* Cpn60 targets a relict organelle. *Current Genetics*, 44(3), pp.138-147.
- Simpson, A.G.B., Inagaki, Y. & Roger, A.J., 2006. Comprehensive multigene phylogenies of excavate protists reveal the evolutionary positions of "primitive" eukaryotes. *Molecular Biology and Evolution*, 23(3), pp.615-625.
- Simpson, A.G.B. et al., 2002. Evolutionary history of "early-diverging" eukaryotes: the excavate taxon *Carpodomonas* is a close relative of *Giardia*. *Molecular Biology and Evolution*, 19(10), pp.1782-1791.
- Stechmann, A. et al., 2008. Organelles in *Blastocystis* that blur the distinction between mitochondria and hydrogenosomes. *Current Biology: CB*, 18(8), pp.580-585.
- Sunila, I., Hamilton, R.M. & Dungan, C.F., 2001. Ultrastructural characteristics of the in vitro cell cycle of the protozoan pathogen of oysters, *Perkinsus marinus*. *The Journal of Eukaryotic Microbiology*, 48(3), pp.348-361.

- Tovar, J, Fischer, A. & Clark, C G, 1999. The mitosome, a novel organelle related to mitochondria in the amitochondrial parasite *Entamoeba histolytica*. *Molecular Microbiology*, 32(5), pp.1013-1021.
- Tovar, J. et al., 2003. Mitochondrial remnant organelles of *Giardia* function in iron-sulphur protein maturation. *Nature*, 426(6963), pp.172-176.
- Tsaousis, A.D. et al., 2008. A novel route for ATP acquisition by the remnant mitochondria of *Encephalitozoon cuniculi*. *Nature*, 453(7194), pp.553-556.
- Yarlett, N., Lloyd, D. & Williams, A.G., 1982. Respiration of the rumen ciliate *Dasytricha ruminantium* Schuberg. *The Biochemical Journal*, 206(2), pp.259-266.

10.2 Appendices for Chapter 2: Methods

10.2.1 Appendix 2 Suppliers of reagents used in this study

A short list has been composed here showing the suppliers for the reagents used in this study, most are well known sources, but some specific items, such as the Optiprep Iodixanol solution were harder to source, and their inclusion in this list might prove useful to anyone repeating this work.

| Reagent | Supplier |
|---|-------------------|
| 5-Sulfosalicylic acid | Fisher Scientific |
| AccuGel 40%(w/v) 29:1 Acrylamide:Bis-acrylamide | Geneflow |
| Agarose, Type I | Calbiochem |
| Amicon ultra-15 centrifugal filter unit, 5kDa | Sigma-Aldrich |
| Amintra Protein A resin | Expedeon |
| Ammonium iron (II) sulfate | Fisher Scientific |
| Ammonium persulfate | Fisher Scientific |
| Anti histidine-tagged protein Mouse Antibody | Calbiochem |
| ASB-14-4 | Calbiochem |
| BIS-Tris, ULTROL | Calbiochem |
| CHAPS | Calbiochem |
| Citric acid, monohydrate | Calbiochem |
| Coomassie Brilliant Blue G250 | Fisher Scientific |
| D-(+)-Maltose monohydrate | Fisher Scientific |
| Digitonin, High purity | Calbiochem |
| DL-Dithiothreitol | Sigma-Aldrich |
| DSP (dithiobis (succinimidyl propionate)) | Thermo scientific |
| DTSSP (3,3' Dithiobis (sulfosuccinimidyl propionate)) | Thermo scientific |
| Formaldehyde | Calbiochem |
| G418 sulfate | Acros Organics |
| Glutaraldehyde solution, 25% | Sigma-Aldrich |
| Glutaraldehyde solution, 50% | Fisher Scientific |
| Glycerol, 100% | Fisher Scientific |
| Guanidine hydrochloride | Fluka |
| HEPES, free acid | Melford |
| His-Tag antibody HRP conjugate | Qiagen |
| Horse Serum, Heat inactivated | Gibco |
| Imidazole | Fisher Scientific |
| Instant Blue | Expedeon |
| IPTG | Fisher Scientific |
| Kanamycin sulfate | Calbiochem |

| | |
|--|----------------------|
| L-Ascorbic acid | Fisher Scientific |
| L-Cysteine, free base | Calbiochem |
| MOPS, free acid, ULTROL | Calbiochem |
| N N' methylenebiacrylamide | Fisher Scientific |
| Ni-NTA His-Bind Superflow beads | Novagen |
| OptiPrep | Axis-Shield |
| OptiPrep density gradient medium | Sigma-Aldrich |
| Paraformaldehyde | Fisher Scientific |
| Penicillin Streptomycin | Gibco |
| Phenylmethylsulfonyl Fluoride | Calbiochem |
| Poncea S | Fisher Scientific |
| Potassium acetate | Sigma-Aldrich |
| Potassium chloride | Sigma-Aldrich |
| Potassium hydroxide | Fisher Scientific |
| Potassium phosphate, dibasic | Calbiochem |
| Potassium phosphate, monobasic | Calbiochem |
| PVDF, Immobilon-P | Millipore |
| Silane-Prep slides | Sigma-Aldrich |
| Sodium Dodecyl Sulfate | National diagnostics |
| Sodium pyruvate | Sigma-Aldrich |
| SuperSignal West Pico Chemiluminescent Substrate | Thermo scientific |
| TEMED | Fisher Scientific |
| Triton X-100, 10% | Calbiochem |
| Triton X-100, 100% | Sigma-Aldrich |
| Trypsin inhibitor, Type II-S: Soybean | Sigma-Aldrich |
| Tryptone | Oxoid |
| TWEEN 20 detergent | Calbiochem |
| U-tube concentrators | Novagen |
| Yeast extract | Oxoid |

Table 10.2 Suppliers of selected materials used in this study

Selected materials which were used to support this investigation have been included in a concise table above. This table aims to address the sources of materials not specifically stated in the body text. Enzymes, primers, antibodies are listed by supplier in text or relevant tables in the methods section.

10.2.2 Appendix 3 pTagVag2 vector, and generation of Tv. Tim44 containing construct

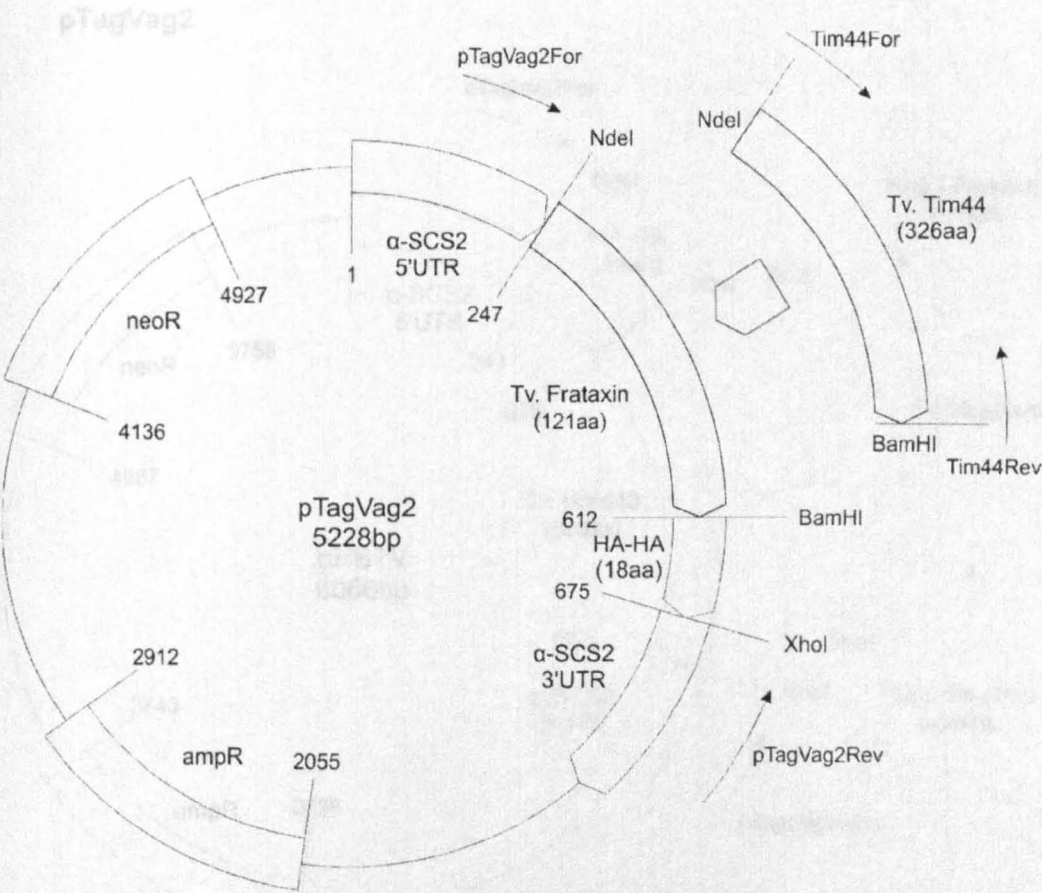


Figure 10.1.2 Basic map of pTagVag2, and the manipulations performed to produce the Tim44 homologue containing construct

The pTagVag2 vector first produced by P. Dolezal was re-used in this investigation to generate HA-tagged frataxin expressing transformants. This template vector was redeveloped to express the *T. vaginalis* Tim44 homologue, by restriction and re-ligation into the insert site using PCR products obtained from gDNA using the Tim44 primers. Insertion of the fragment utilizes the same restriction enzymes employed in the insertion of frataxin, whereby the start codon is synthesized from the NdeI restriction site, sequencing of the insert was performed using two sequencing primers outside of the protein coding region.

10.2.3 Appendix 4 Generation of the β -barrel containing expression vectors from pTagVag2

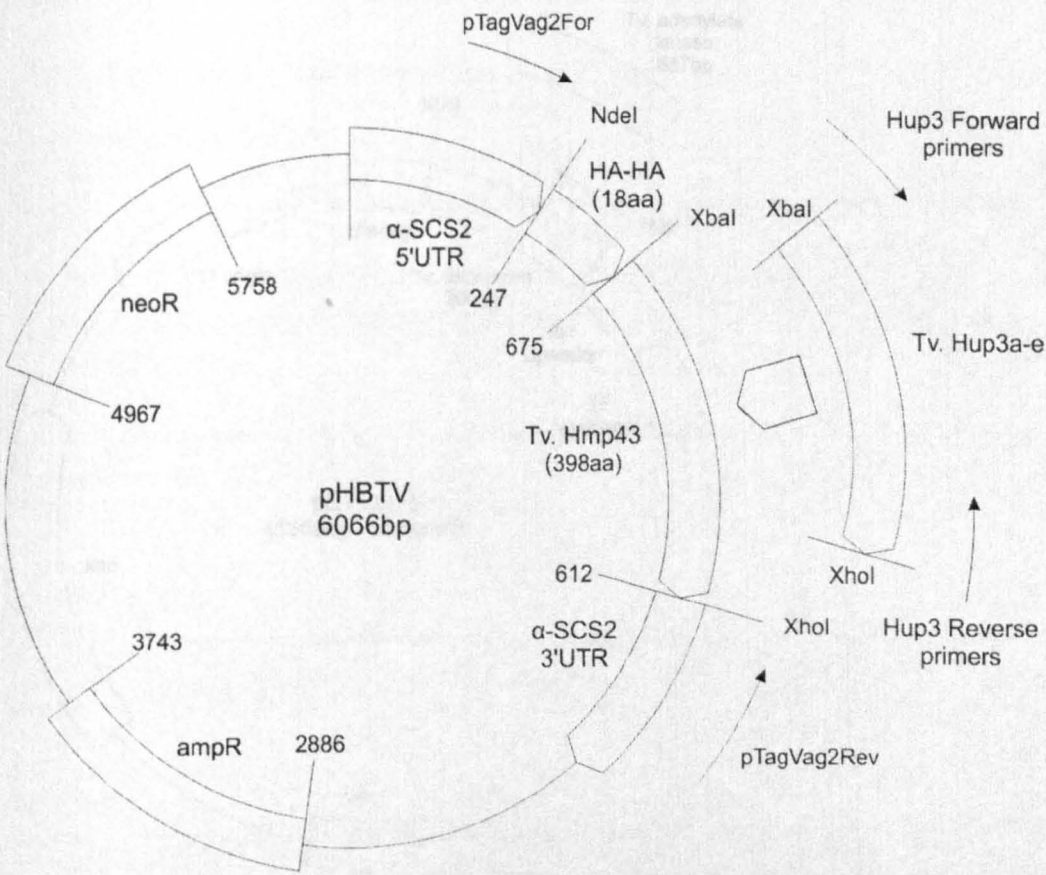


Figure 10.1.3 A simple map of the β -barrel expression constructs modeled from pTagVag2

The pTagVag2 vector was redeveloped to express candidate β -barrel proteins, a template plasmid was produced with Hmp43, with the double HA-tag moved N-terminally of the inserted sequence. The start codon was synthesized from the NdeI restriction site. This template was used to produce subsequent Hup3 containing vectors, by restriction and re-ligation of gDNA PCR products produced with Hup3 primers. The pTagVag2 sequence primers were used to verify sequence fidelity, and are position outside of the insert region.

10.2.4 Appendix 5 Generation of constructs expressing *Tv. ferredoxin* and adenylate kinase

10.3 Appendices for Chapter 6: Bioinformatics

10.3.1 Appendix 6 Ton20 like proteins

5. coreview20 query protein

GI285814534

5000 aa

1000 aa

2000 aa

3000 aa

4000 aa

5000 aa

6000 aa

7000 aa

8000 aa

9000 aa

10000 aa

11000 aa

12000 aa

13000 aa

14000 aa

15000 aa

16000 aa

17000 aa

18000 aa

19000 aa

20000 aa

21000 aa

22000 aa

23000 aa

24000 aa

25000 aa

26000 aa

27000 aa

28000 aa

29000 aa

30000 aa

31000 aa

32000 aa

33000 aa

34000 aa

35000 aa

36000 aa

37000 aa

38000 aa

39000 aa

40000 aa

41000 aa

42000 aa

43000 aa

44000 aa

45000 aa

46000 aa

47000 aa

48000 aa

49000 aa

50000 aa

51000 aa

52000 aa

53000 aa

54000 aa

55000 aa

56000 aa

57000 aa

58000 aa

59000 aa

60000 aa

61000 aa

62000 aa

63000 aa

64000 aa

65000 aa

66000 aa

67000 aa

68000 aa

69000 aa

70000 aa

71000 aa

72000 aa

73000 aa

74000 aa

75000 aa

76000 aa

77000 aa

78000 aa

79000 aa

80000 aa

81000 aa

82000 aa

83000 aa

84000 aa

85000 aa

86000 aa

87000 aa

88000 aa

89000 aa

90000 aa

91000 aa

92000 aa

93000 aa

94000 aa

95000 aa

96000 aa

97000 aa

98000 aa

99000 aa

100000 aa

101000 aa

102000 aa

103000 aa

104000 aa

105000 aa

106000 aa

107000 aa

108000 aa

109000 aa

110000 aa

111000 aa

112000 aa

113000 aa

114000 aa

115000 aa

116000 aa

117000 aa

118000 aa

119000 aa

120000 aa

121000 aa

122000 aa

123000 aa

124000 aa

125000 aa

126000 aa

127000 aa

128000 aa

129000 aa

130000 aa

131000 aa

132000 aa

133000 aa

134000 aa

135000 aa

136000 aa

137000 aa

138000 aa

139000 aa

140000 aa

141000 aa

142000 aa

143000 aa

144000 aa

145000 aa

146000 aa

147000 aa

148000 aa

149000 aa

150000 aa

151000 aa

152000 aa

153000 aa

154000 aa

155000 aa

156000 aa

157000 aa

158000 aa

159000 aa

160000 aa

161000 aa

162000 aa

163000 aa

164000 aa

165000 aa

166000 aa

167000 aa

168000 aa

169000 aa

170000 aa

171000 aa

172000 aa

173000 aa

174000 aa

175000 aa

176000 aa

177000 aa

178000 aa

179000 aa

180000 aa

181000 aa

182000 aa

183000 aa

184000 aa

185000 aa

186000 aa

187000 aa

188000 aa

189000 aa

190000 aa

191000 aa

192000 aa

193000 aa

194000 aa

195000 aa

196000 aa

197000 aa

198000 aa

199000 aa

200000 aa

201000 aa

202000 aa

203000 aa

204000 aa

205000 aa

206000 aa

207000 aa

208000 aa

209000 aa

210000 aa

211000 aa

212000 aa

213000 aa

214000 aa

215000 aa

216000 aa

217000 aa

218000 aa

219000 aa

220000 aa

221000 aa

222000 aa

223000 aa

224000 aa

225000 aa

226000 aa

227000 aa

228000 aa

229000 aa

230000 aa

231000 aa

232000 aa

233000 aa

234000 aa

235000 aa

236000 aa

237000 aa

238000 aa

239000 aa

240000 aa

241000 aa

242000 aa

243000 aa

244000 aa

245000 aa

246000 aa

247000 aa

248000 aa

249000 aa

250000 aa

251000 aa

252000 aa

253000 aa

254000 aa

255000 aa

256000 aa

257000 aa

258000 aa

259000 aa

260000 aa

261000 aa

262000 aa

263000 aa

264000 aa

265000 aa

266000 aa

267000 aa

268000 aa

269000 aa

270000 aa

271000 aa

272000 aa

273000 aa

274000 aa

275000 aa

276000 aa

277000 aa

278000 aa

279000 aa

280000 aa

281000 aa

282000 aa

283000 aa

284000 aa

285000 aa

286000 aa

287000 aa

288000 aa

289000 aa

290000 aa

291000 aa

292000 aa

293000 aa

294000 aa

295000 aa

296000 aa

297000 aa

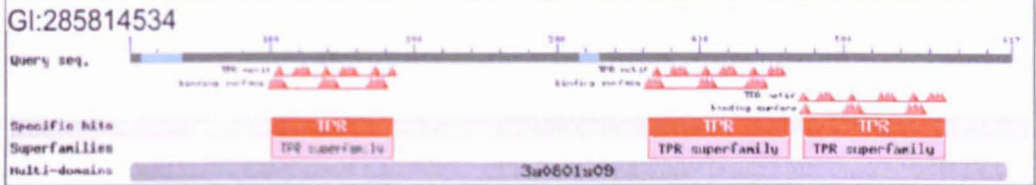
298000 aa

299000 aa

10.3 Appendices for Chapter 5: Bioinformatics

10.3.1 Appendix 6 Tom70 like proteins in *T. vaginalis*

S.cerevisiae query protein



T.vaginalis proteins



Figure10.3.1 Domain organization of Tom70 like proteins in *T. vaginalis*

The BLAST results for the TOM complex receptor Tom70 yielded multiple results, which warranted further investigation, however many of the result proteins were of truncated size, the remaining sequences were examined in CDD (Marchler-Bauer et al. 2011) and the graphical result of its domain prediction are shown above. Whilst sequence similarity to Tom70 is very good in these sequences the domain architecture and arrangement are significantly different to the original mitochondrial sequence. The repeated domain is likely the reason for the unusually high homology.

10.3.2 Appendix 7 transmembrane prediction of Hup3 proteins

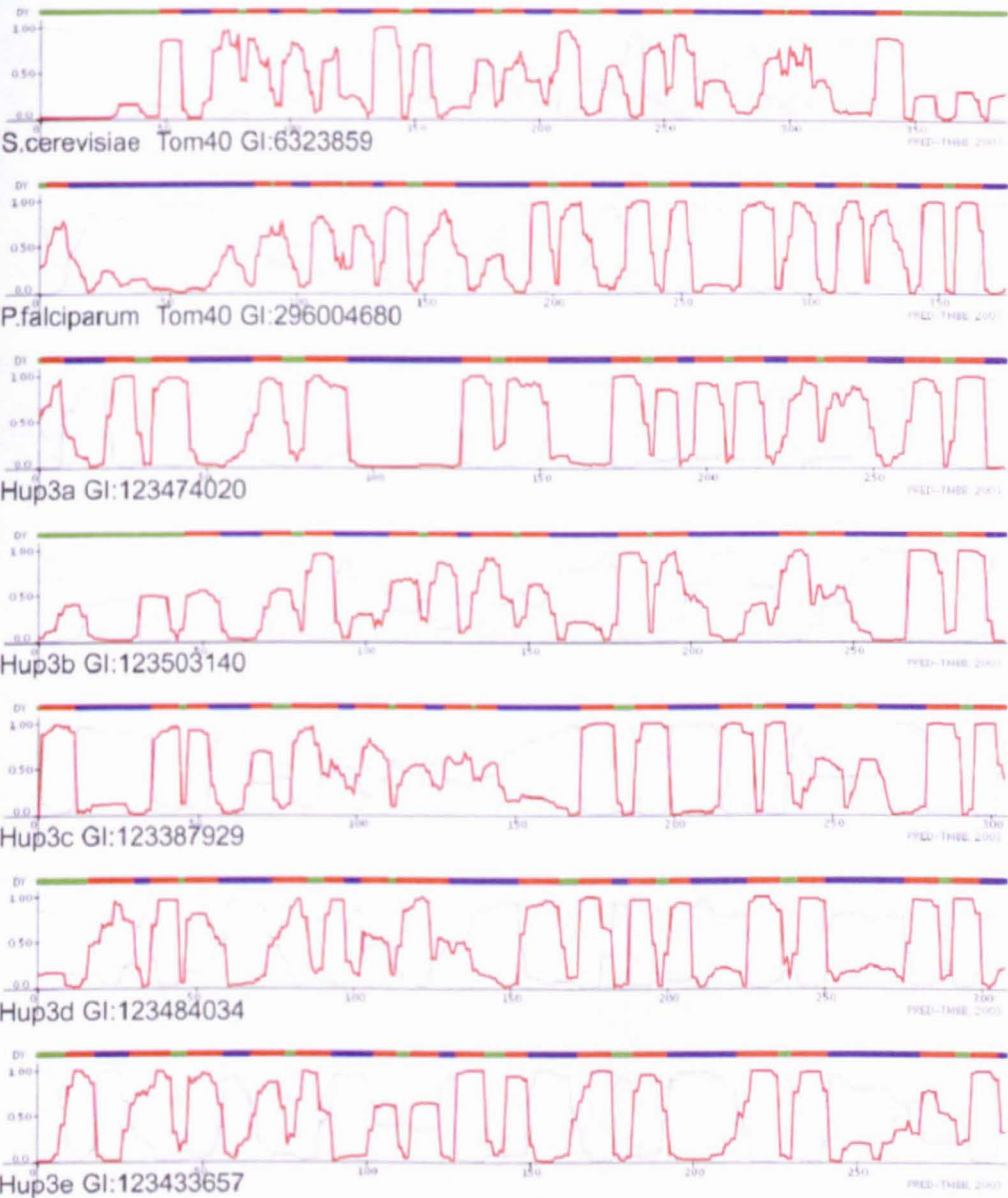


Figure 10.3.2 Predicted β -barrel strand prediction in the Hup3 family proteins

The secondary structure of the *T. vaginalis* Hup3 proteins was predicted using the Pred-TMBB tool, which uses a HMM algorithm to predict the positions of transmembrane strands of β -barrel proteins. The results shown above indicate the positions of predicted strands (as plateaus). consistent features are seen between Hup3 proteins, and the C-terminal region of *S. cerevisiae* Tom40.

10.3.3 Appendix 8 transmembrane prediction of Hmp43

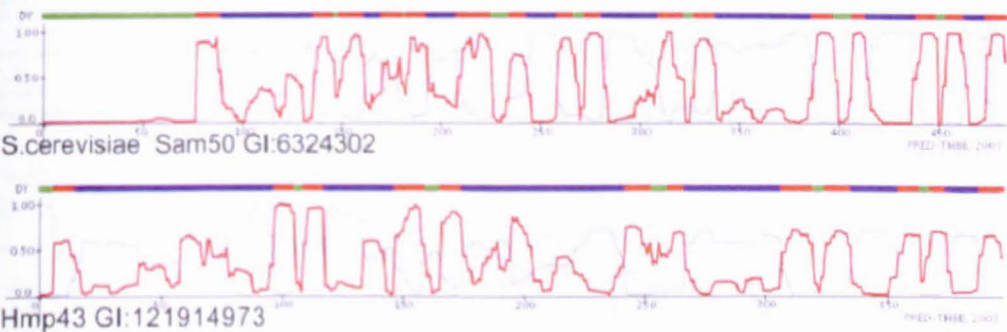


Figure 10.3.3 Predicted transmembrane structure of Hmp43

The Pred-TMBB tool was used to analyze the primary sequence of *T. vaginalis* Hmp43, and *S. cerevisiae* Sam50. The results are shown above. Transmembrane β -strands are presented a flat topped plateaus, when examining the relation of the two proteins above, consideration must be given to the differences in length, however a clearly conserved C-terminal structure is seen shared between both proteins.

10.3.4 Appendix 9 PAM homologues in *T.vaginalis*

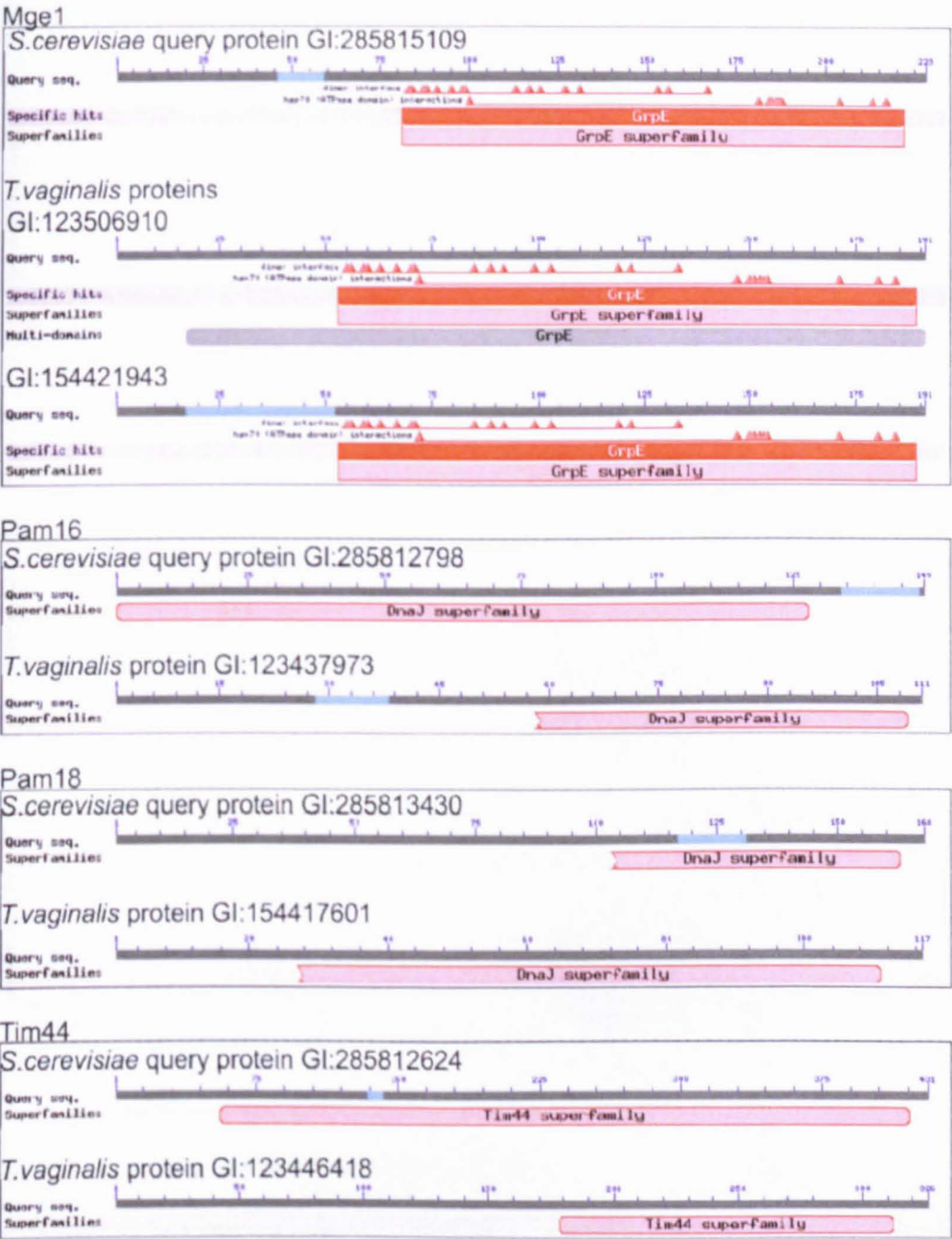


Figure 10.3.4 Similarity of *T. vaginalis* and *S. cerevisiae* PAM components as visualised by CDD

To verify the similarity seen in BLAST between *T. vaginalis* and *S. cerevisiae* PAM components, domain architecture was examined in CDD. Domain position and organisation are favourably similar, and indicate that the observed homology represents true conservation of sequence architecture.

10.3.5 Appendix 10 TopPred topology prediction of *S. cerevisiae* Tim17 superfamily proteins

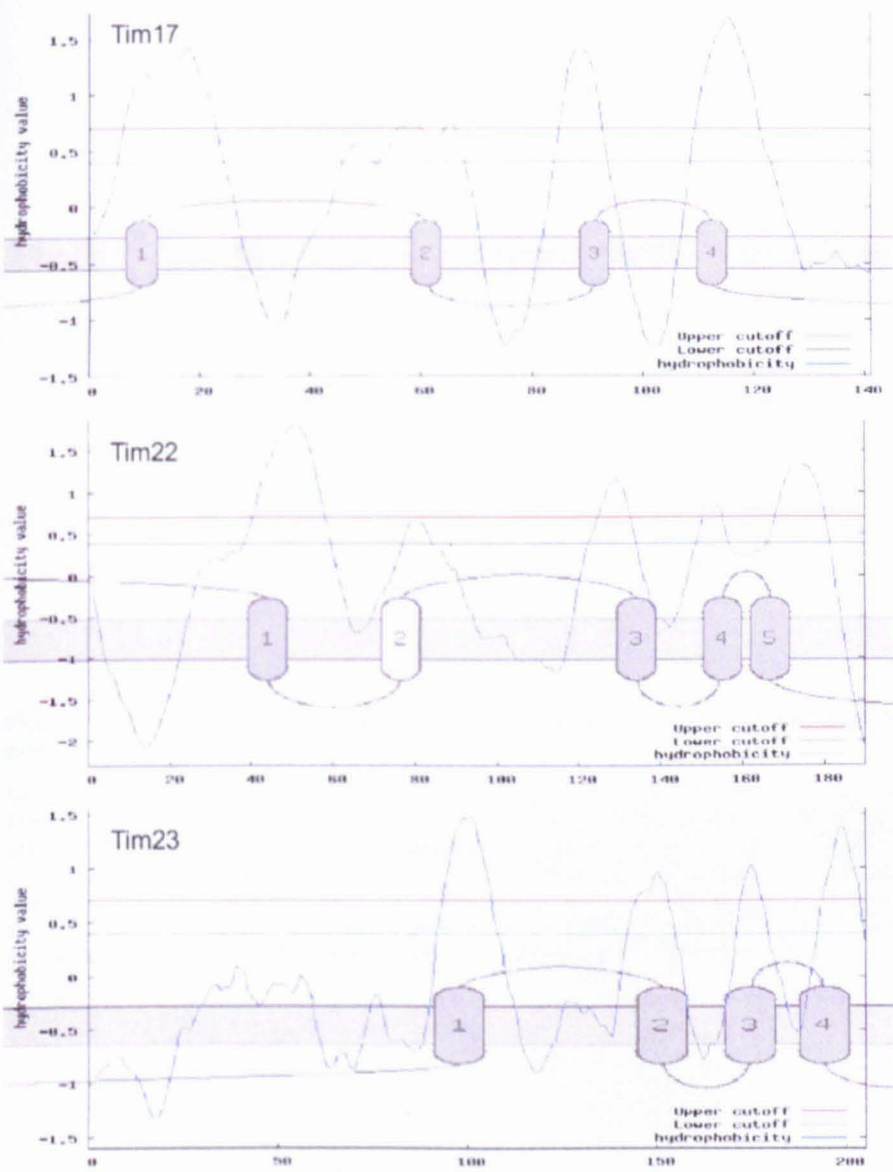


Figure10.3.5 Topology prediction of Tim17 superfamily proteins from *S. cerevisiae*

The membrane protein topology prediction program TopPred was used to predict the topology of *S.cerevisiae* Tim17,22,23 for comparison against the *T. vaginalis* Phat proteins. Putative (green) and certain cutoffs (red) were used with values of 0.3 and 0.7. A window size of 5 was used to detect the short loops present within this tightly packed domain, and the Kyte and Doolittle hydrophobicity scale was used to rate the hydrophobicity of the primary sequence. The Tim17 domain is seen as a collection of 4 C-terminal transmembrane helices. In Tim22 an additional helix of 'predicted' in the N-terminal region, however, this protein is not known to have any additional helices out of the core Tim17 domain, and likely a program artifact.

10.3.6 Appendix 11 Detection of multiple Hsp70 homologues

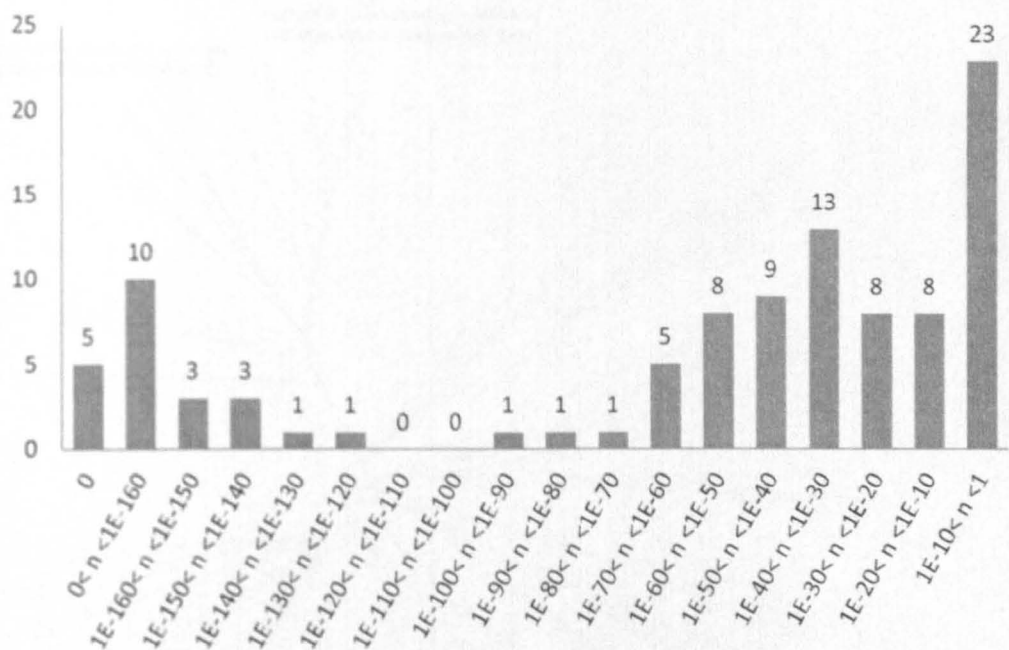


Figure 10.3.6 Distribution of sequence hits amongst *T. vaginalis* proteins homologous to *S. cerevisiae* mtHsp70

Using the BLAST algorithm under stringent conditions (BLOSUM 80 matrix) to find highly homologous sequences to the *S. cerevisiae* mtHsp70 protein, recovers 100 *T. vaginalis* proteins with an E values of <1. Five of these proteins are identical to portions of the query sequence, however the majority of the proteins span a wide range of values. This graph displays the distribution of BLAST E values for the recovered proteins. If all copies of the *T. vaginalis* mtHsp70 diverged uniformly from the query sequence then the distribution of proteins would peak and subside around a common point. This graph illustrates an irregular population of divergent proteins, with complex phylogeny. The analysis of these results are shown in Appendices 10.3.7, 10.3.8.

Figure 10.3.7 Distribution of sequence hits amongst *T. vaginalis* proteins homologous to *S. cerevisiae* mtHsp70

Figure 10.3.8 Distribution of sequence hits amongst *T. vaginalis* proteins homologous to *S. cerevisiae* mtHsp70

Figure 10.3.9 Distribution of sequence hits amongst *T. vaginalis* proteins homologous to *S. cerevisiae* mtHsp70

Figure 10.3.10 Distribution of sequence hits amongst *T. vaginalis* proteins homologous to *S. cerevisiae* mtHsp70

Figure 10.3.11 Distribution of sequence hits amongst *T. vaginalis* proteins homologous to *S. cerevisiae* mtHsp70

Figure 10.3.12 Distribution of sequence hits amongst *T. vaginalis* proteins homologous to *S. cerevisiae* mtHsp70

Figure 10.3.13 Distribution of sequence hits amongst *T. vaginalis* proteins homologous to *S. cerevisiae* mtHsp70

Figure 10.3.14 Distribution of sequence hits amongst *T. vaginalis* proteins homologous to *S. cerevisiae* mtHsp70

Figure 10.3.15 Distribution of sequence hits amongst *T. vaginalis* proteins homologous to *S. cerevisiae* mtHsp70

Figure 10.3.16 Distribution of sequence hits amongst *T. vaginalis* proteins homologous to *S. cerevisiae* mtHsp70

Figure 10.3.17 Distribution of sequence hits amongst *T. vaginalis* proteins homologous to *S. cerevisiae* mtHsp70

Figure 10.3.18 Distribution of sequence hits amongst *T. vaginalis* proteins homologous to *S. cerevisiae* mtHsp70

Figure 10.3.19 Distribution of sequence hits amongst *T. vaginalis* proteins homologous to *S. cerevisiae* mtHsp70

Figure 10.3.20 Distribution of sequence hits amongst *T. vaginalis* proteins homologous to *S. cerevisiae* mtHsp70

Figure 10.3.21 Distribution of sequence hits amongst *T. vaginalis* proteins homologous to *S. cerevisiae* mtHsp70

Figure 10.3.22 Distribution of sequence hits amongst *T. vaginalis* proteins homologous to *S. cerevisiae* mtHsp70

Figure 10.3.23 Distribution of sequence hits amongst *T. vaginalis* proteins homologous to *S. cerevisiae* mtHsp70

Figure 10.3.24 Distribution of sequence hits amongst *T. vaginalis* proteins homologous to *S. cerevisiae* mtHsp70

IMAGING SERVICES NORTH

Boston Spa, Wetherby
West Yorkshire, LS23 7BQ
www.bl.uk

TEXT BOUND CLOSE TO
THE SPINE IN THE
ORIGINAL THESIS

10.3.7 Appendix 12 Hsp70-like protein families in *T. vaginalis*

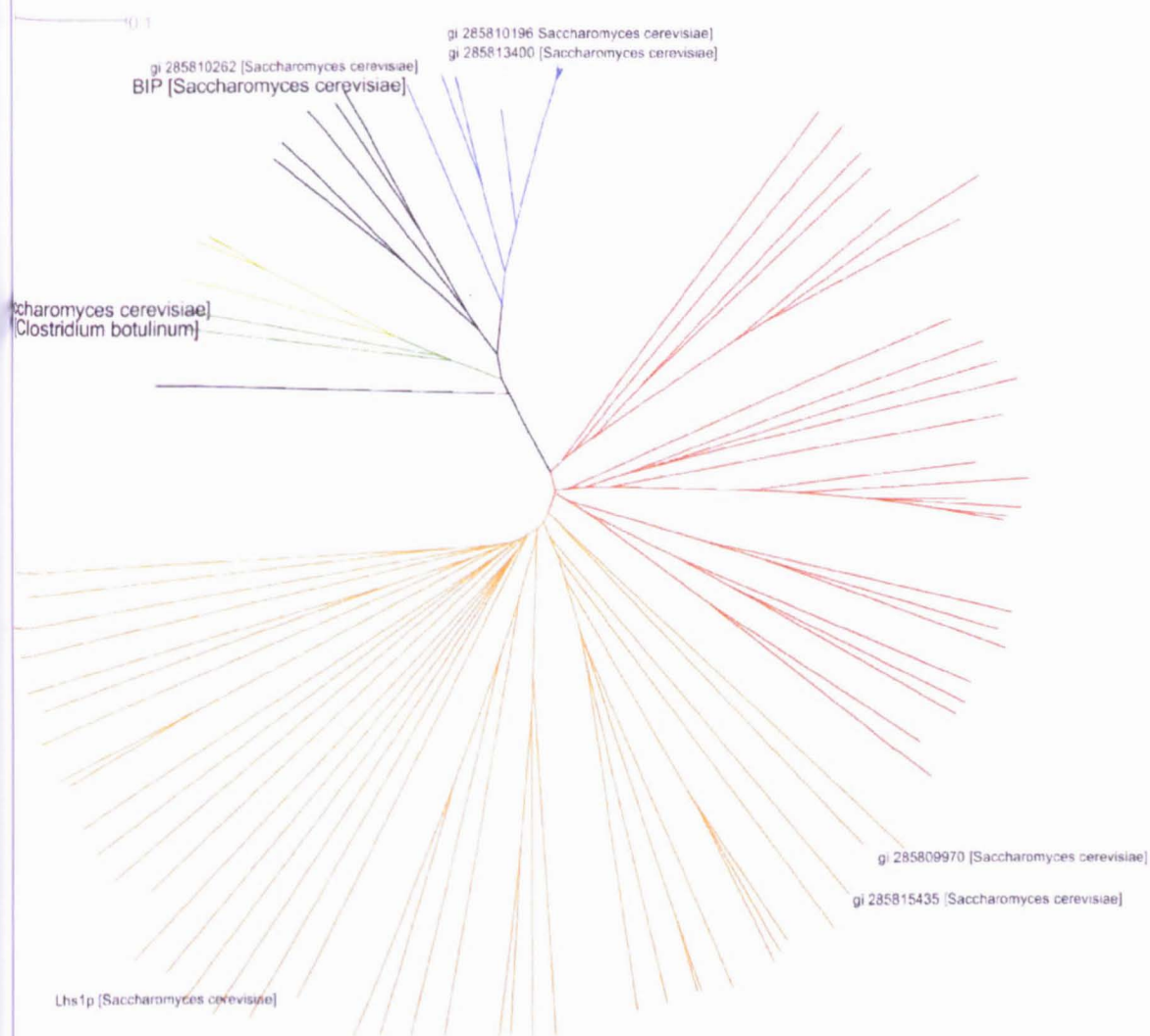


Table10.3 Identification of a mtHsp70 subclade from other dnaK homologues

Within *S. cerevisiae* there are many sequences with a shared dnaK core domain, which function in diverse roles in protein chaperoning and trafficking in the cell. The large number of returned hits in *T. vaginalis* is not a result of selective expansion of the mtHsp70 subfamily of these proteins. This was determined by collecting the different dnaK containing proteins of *S. cerevisiae*, and an ancestral dnaK sequence from clostridium, and using these in an alignment with the *T. vaginalis* results. Using this technique the *T. vaginalis* proteins clade to their closest *S. cerevisiae* homologues. This approach was used to determine the most likely candidates of the mtHsp70 in *T.vaginalis* (green). This approach was also used to identify homologues to the dnaK related protein BIP, an ER chaperone (blue). This determination is useful in the interpretation of the anti-BIP microscopy.

10.3.8 Appendix 13 Identification of mtHsp70 homologues

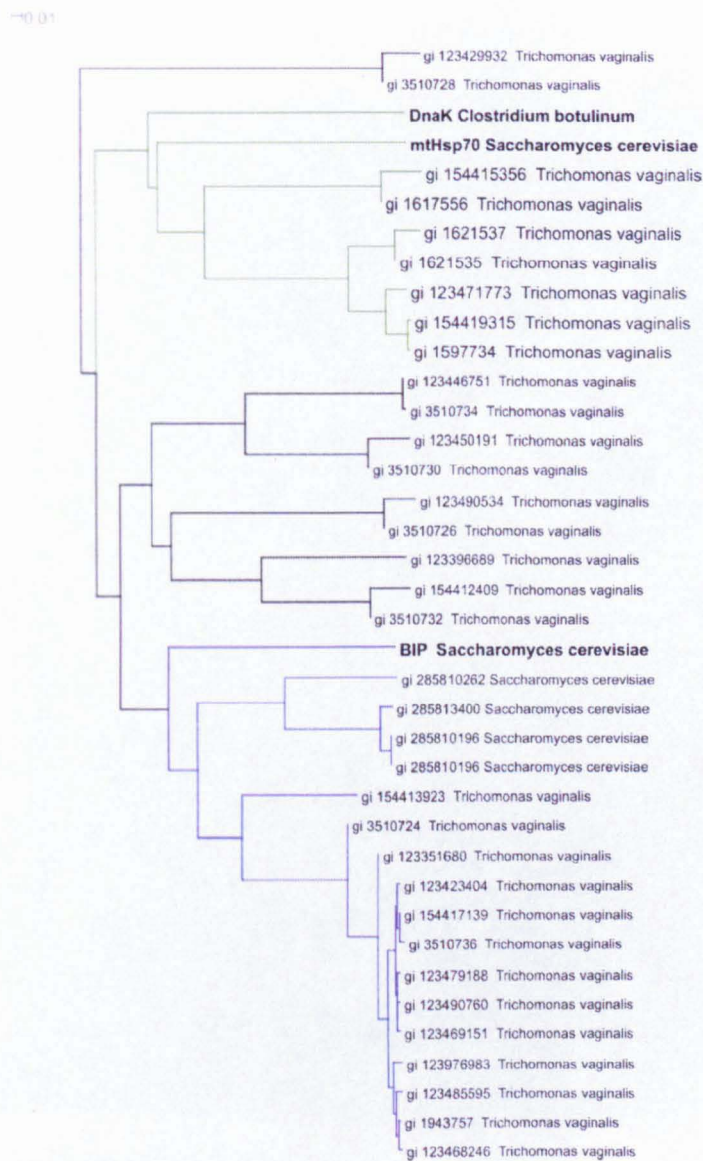


Figure 10.3.7 mtHsp70 and BIP homologues in *T. vaginalis*

The phylogenetic analysis in Appendix 10.3.7 was used to narrow down the dnaK domain containing homologues and arrive at a list of candidates for mtHsp70 and BIP on the basis of cladistic analysis with proteins from *S. cerevisiae*. The subsection of the tree derived in Appendix 10.3.7 is seen above, with the same consistent coloring for mtHsp70 (green), and BIP (blue). *S. cerevisiae* and *C. botulin* proteins are indicated in bold. These data indicate that there are convincing homologues for BIP in *T. vaginalis*, (which is later detected with a discrete localisation pattern in microscopy 6.3.3 p153)

10.4 Appendices for Chapter 6: Microscopy

10.4.1 Appendix 14 Confocal analysis of DAPI stained cells

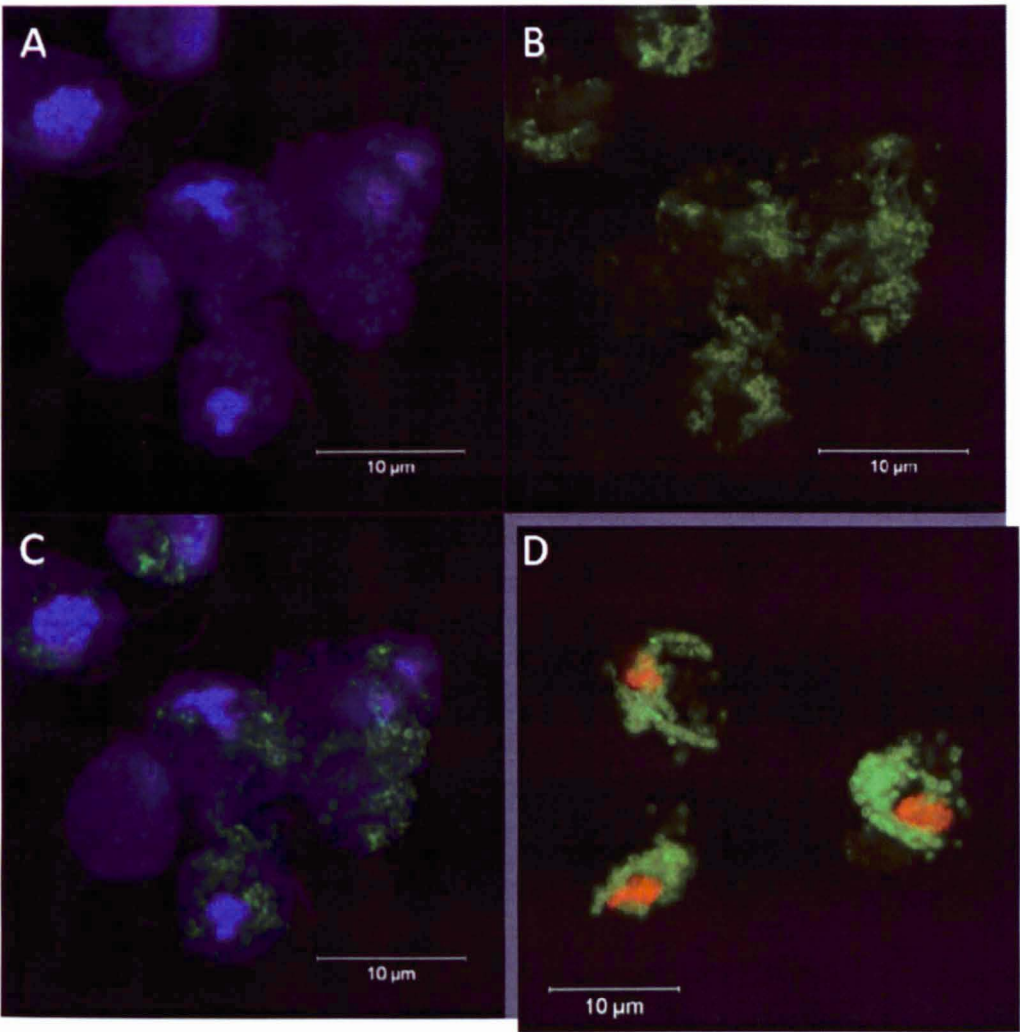


Figure 10.4.1 Confocal analysis of DAPI stained cells

The DAPI stain was further investigated in confocal microscopy, fixed Hup3c transformant cells were treated with DAPI (A) and anti-HA antibody (B). The DAPI channel was illuminated at 405nm and the anti-HA/Alexa448 coupled antibody complex at 488nm. The image was taken with an oil immersed Plan-Apochromat 63x lens, the width of the image area is 33.9μm with a resolution of 0.033 μm/pixel. Light from the DAPI channel was collected using a 410 – 477nm filter, light collection for the Alexa488 antibody was taken at longer wavelengths. These settings effectively eliminated bleed from the DAPI (A) into the Alexa488 channel (B), however diffuse localisation of DAPI is still apparent (A,C). These images are shown against a propidium iodide stained Hup3c cells (D), showing the greater specificity of the Propidium Iodide stain.

The DAPI stain shown in Figure 10.4.1 does demonstrate some interesting features particularly the heterogeneous distribution of signal within the nuclear region, which is not as readily apparent in the PI stained cells. This effect might arise from the emission wavelengths of the two different markers, the considerably longer wavelength of the PI emitted light may effect visible resolution of these structures. However PI does demonstrate a higher specificity for the nucleus as well as a limited background signal, which is more desirable for its use as a nuclear marker.

10.4.2 Appendix 15 Confocal microscopy of PI stained *T. vaginalis* cells

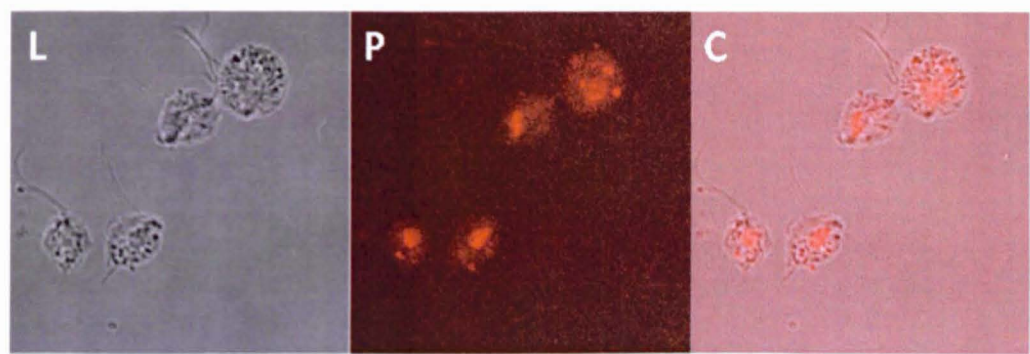


Figure 10.4.2 Confocal microscopy of PI stained *T. vaginalis* cells

The intracellular distribution of the PI nuclear stained was investigated further in confocal microscopy. A transmission image of the cells was taken (L) between 410-470nm using a 405nm laser, the emission of the PI channel was collected between 603-684nm (P). 405nm and 561nm lasers were used to illuminate the sample. The channels are shown combined in third image (C). Large intra-cellular structures of limited internal contrast can be seen in L, which have specifically been labeled with PI. The size and distribution of these structures is fitting with the nucleus.

10.4.3 Appendix 16 A 3D reconstruction of a Golgi apparatus

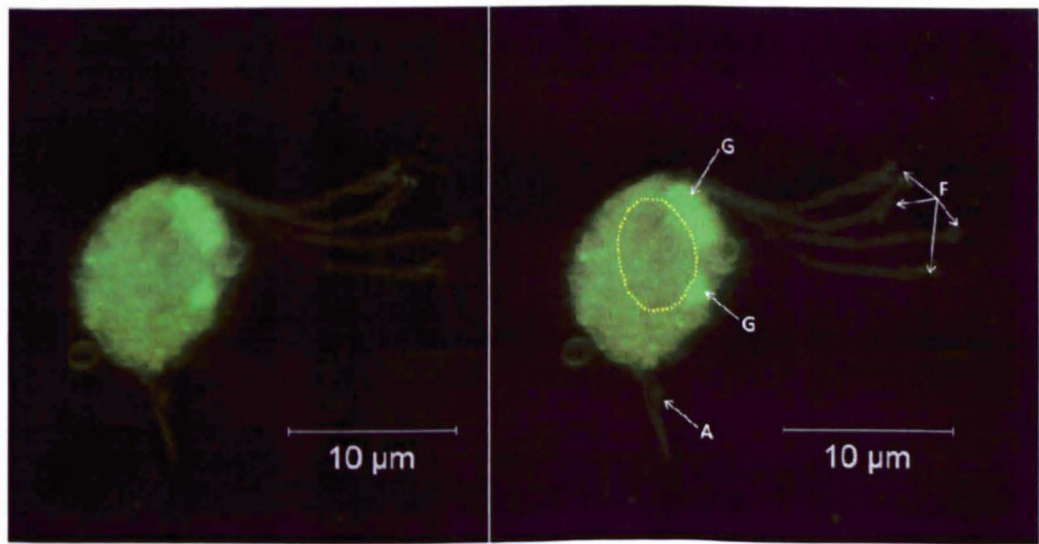


Figure 10.4.3 A 3D reconstruction of a Golgi apparatus

The internal arrangement of a fixed C1 cell was probed with the Golgi marker. In this image a reconstruction of a single cell was performed by combining 22 slices through a volume of 44.90 µm x 44.90 µm x 13.38 µm (z-axis). The Golgi marker was illuminated using a 458nm laser and emitted light collected between 487-597nm. The right hand image is annotated to show Flagella (F), Golgi (G) and Axostyle (A), a signal-intensity void is shown in dotted yellow, indicating the extents of the nucleus.

10.4.4 Appendix 17 Epifluorescent images of Transformant Hup3d

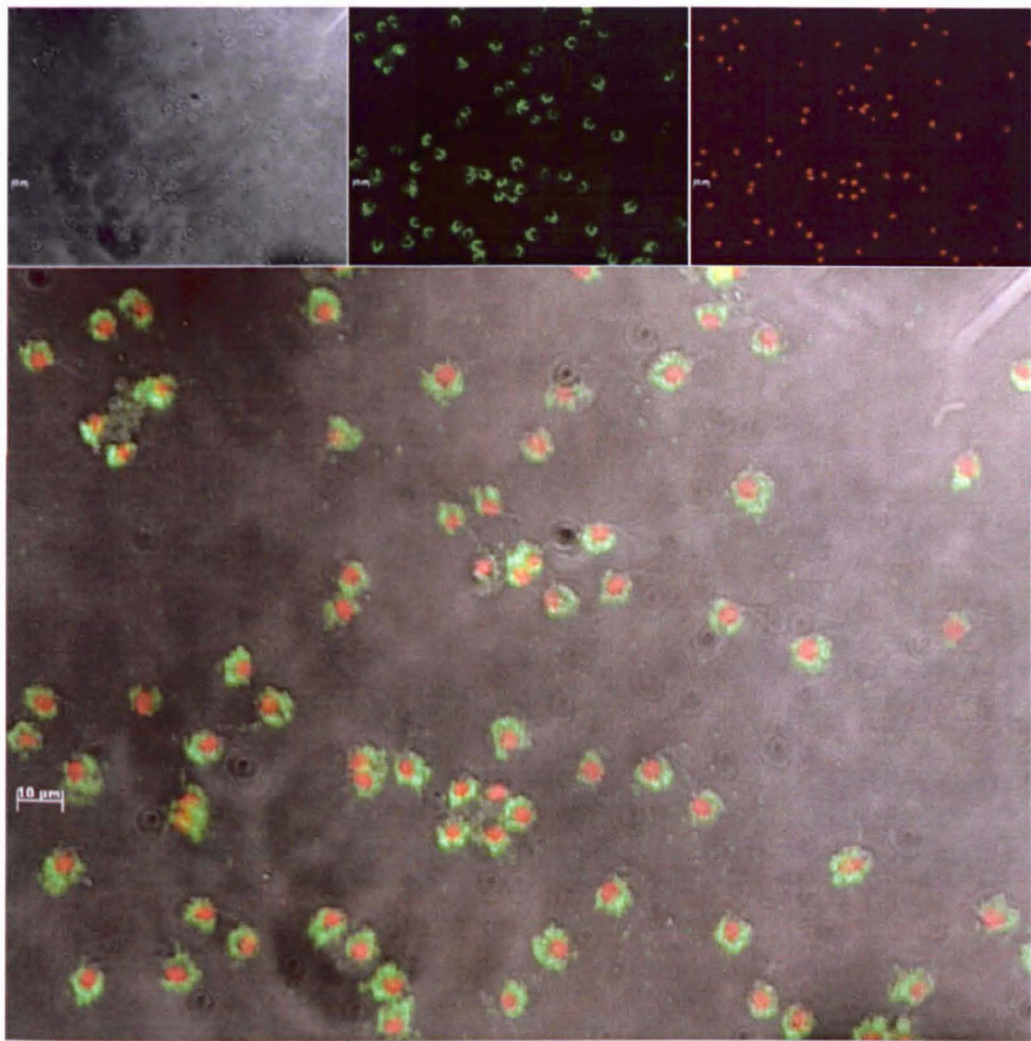


Figure 10.4.4 Epifluorescent images of Hup3d

In the course of optimizing the Propidium nuclear marker (red), images were taken using the epifluorescent microscope of transformant cells where the HA- Hup3d protein was detected using an anti-HA primary/Alexa488 secondary antibody (green). The images show a clear mutual exclusion of the two signal sources. Images were captured using a 63x Oil Immersion lens, with the Alexa488 detected on a FITC channel, and the PI on a Rhodamine channel. Illumination was with a deuterium UV lamp.

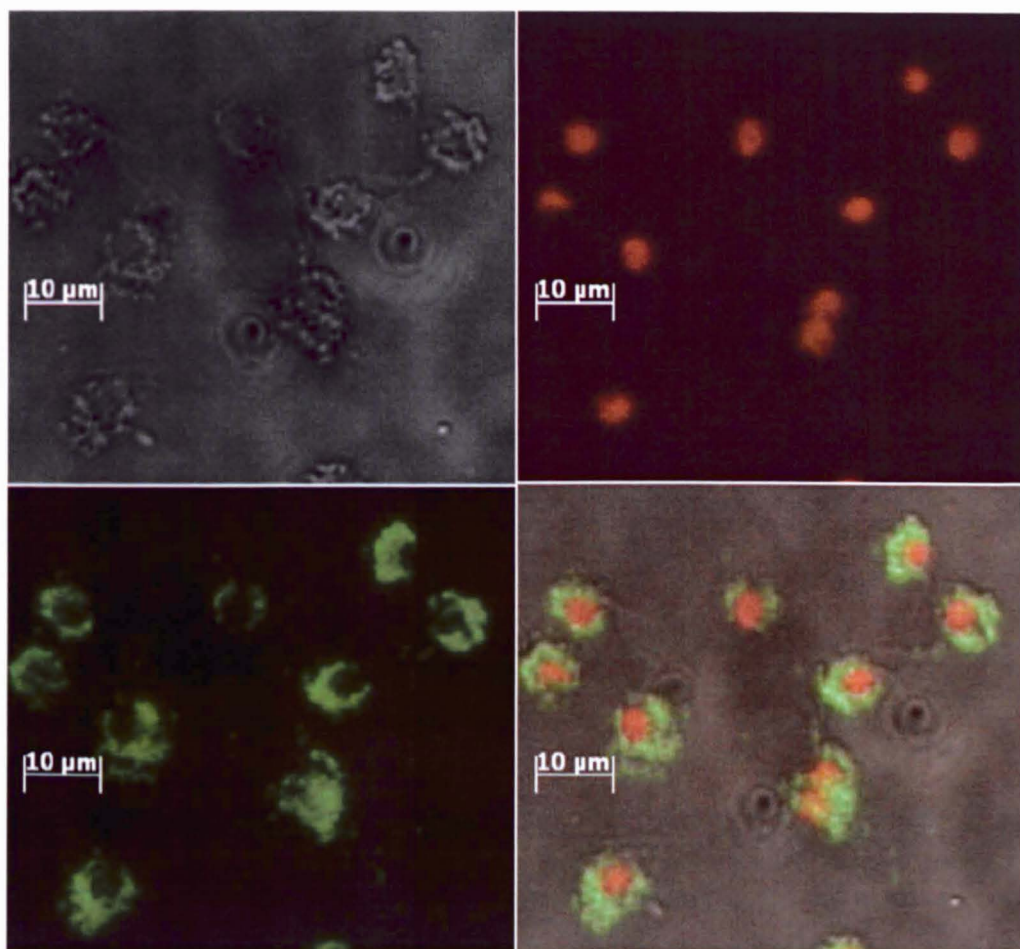


Figure 10.4.5 Epifluorescent images of the localisation of Hup3d

This image is a subset of Figure 10.4.4, showing a close up of the localisation of the transformant protein (green), and nuclear marker PI (red) against a light microscopy image of the fixed cells (grey). There is a clear exclusion of signal between the channels, but further detail is beyond the capabilities of the epifluorescent microscope. For this reason confocal microscopy was chosen to complete higher resolution imaging of transformant cells.

10.4.5 Appendix 18 3D reconstruction of frataxin transformant cells

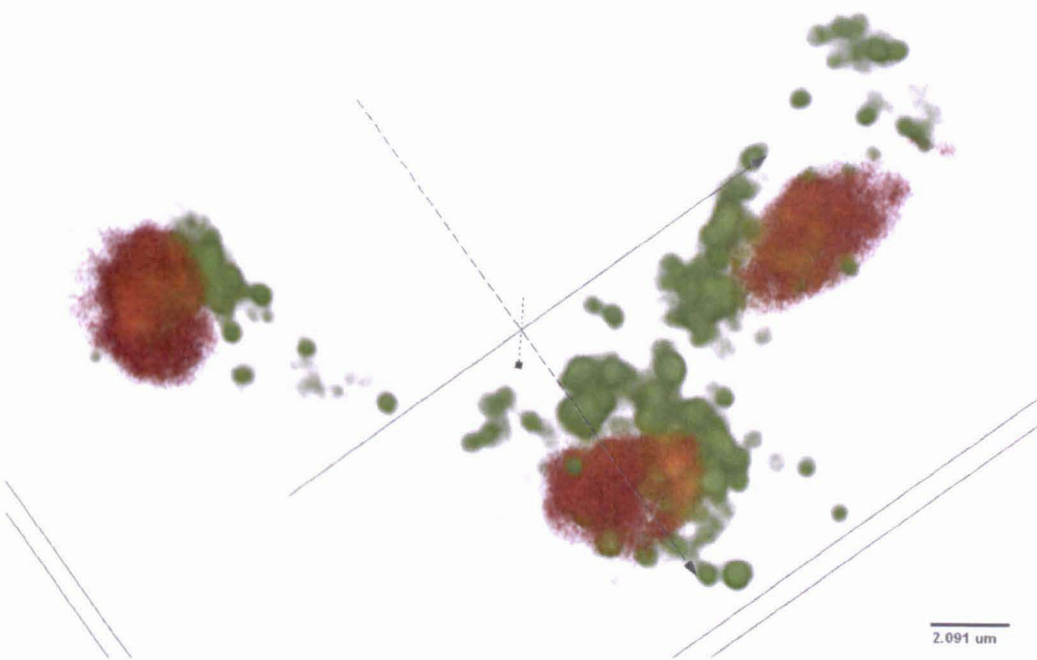


Figure 10.4.6 3D reconstruction of frataxin transformant cells

Z-stack images obtained from confocal microscopy (Figure 6.4.1) were re-processed in the bioView3d program. The program uses a threshold to decide the limits for transparency, and regions with a signal intensity greater than threshold are retained. The reprocessed images are shown here 3D reconstructed to show the ultrastructural arrangements of hydrogenosomes from fixed cells (green), as well as the position of the nucleus indicated by PI marker (red).

10.4.6 Appendix 19 3D reconstruction of Tim44 transformant cells

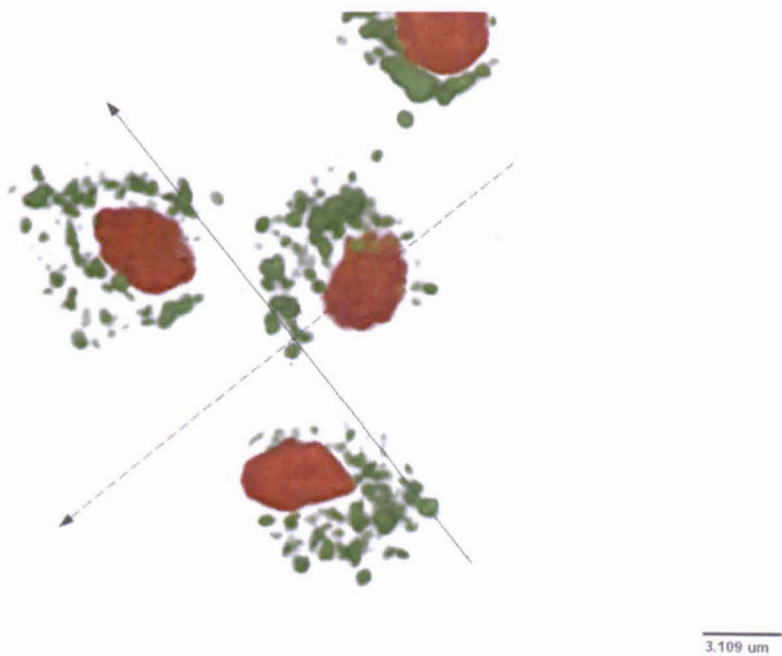


Figure 10.4.7 3D reconstruction of cells from the tim44 transformant

Z-stack image data collected on the tim44 transformant (see Figure 6.4.3) was reconstructed in bioView3d. The image slices were recreated as voxel volume data, and represented here in 3D. The Tim44 containing hydrogenosome structures highlighted here (green) differ slightly to Figure 10.4.6, cell nuclei are also shown in this reconstruction in red.

10.4.7 Appendix 20 Distribution of HA- Hup3a

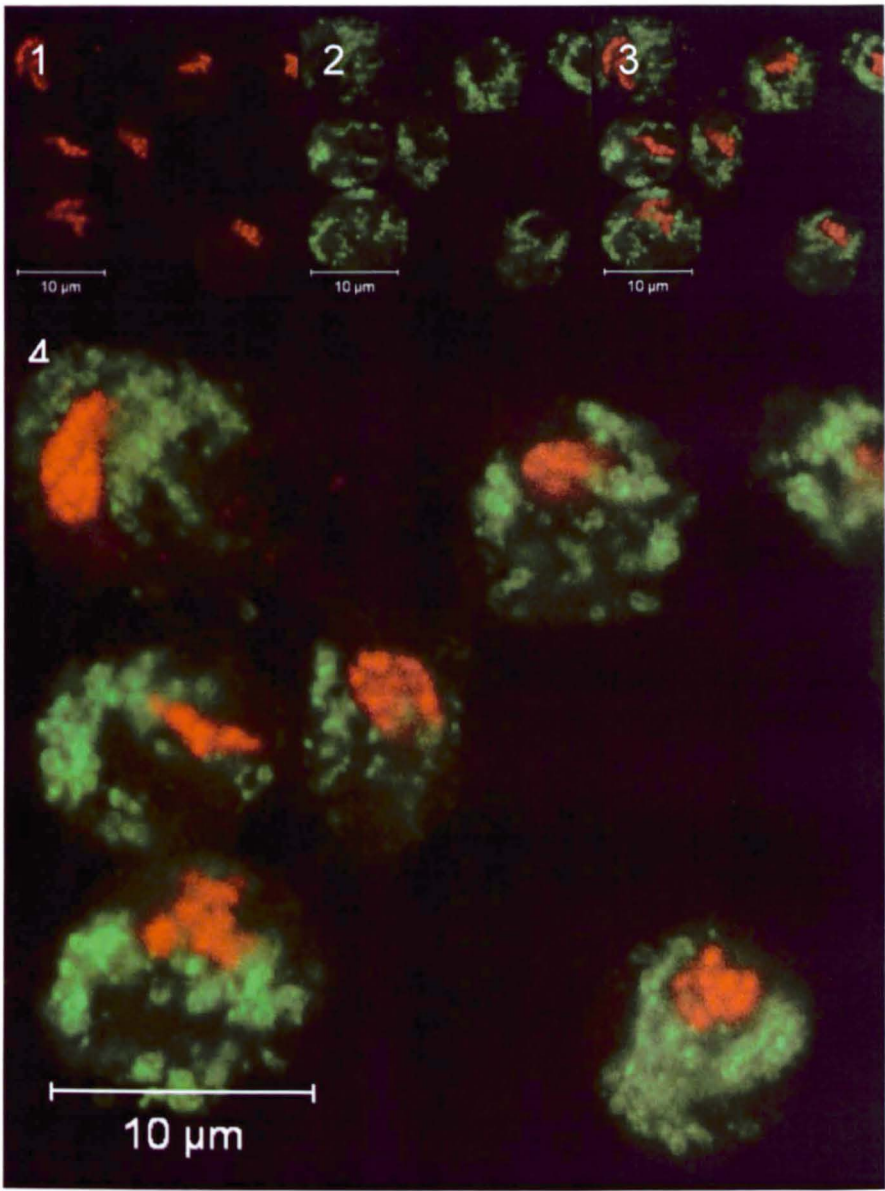


Figure 10.4.8 Distribution of HA-Hup3a

Hup3a expressing transformants were fixed and stained with the nuclear marker PI (1) and fluorescently tagged secondary/anti-HA primary antibody (2). These channels are shown combined in 3,4. Excitation of the PI channel was with a 561nm laser with light collected >600nm. The Alexa488 secondary antibody was excited using a 488nm laser with fluorescent light collected between 493-555nm. The image resolution was captured at 0.066 μm/pixel.

10.4.8 Appendix 21 Intracellular distribution of Hup3c

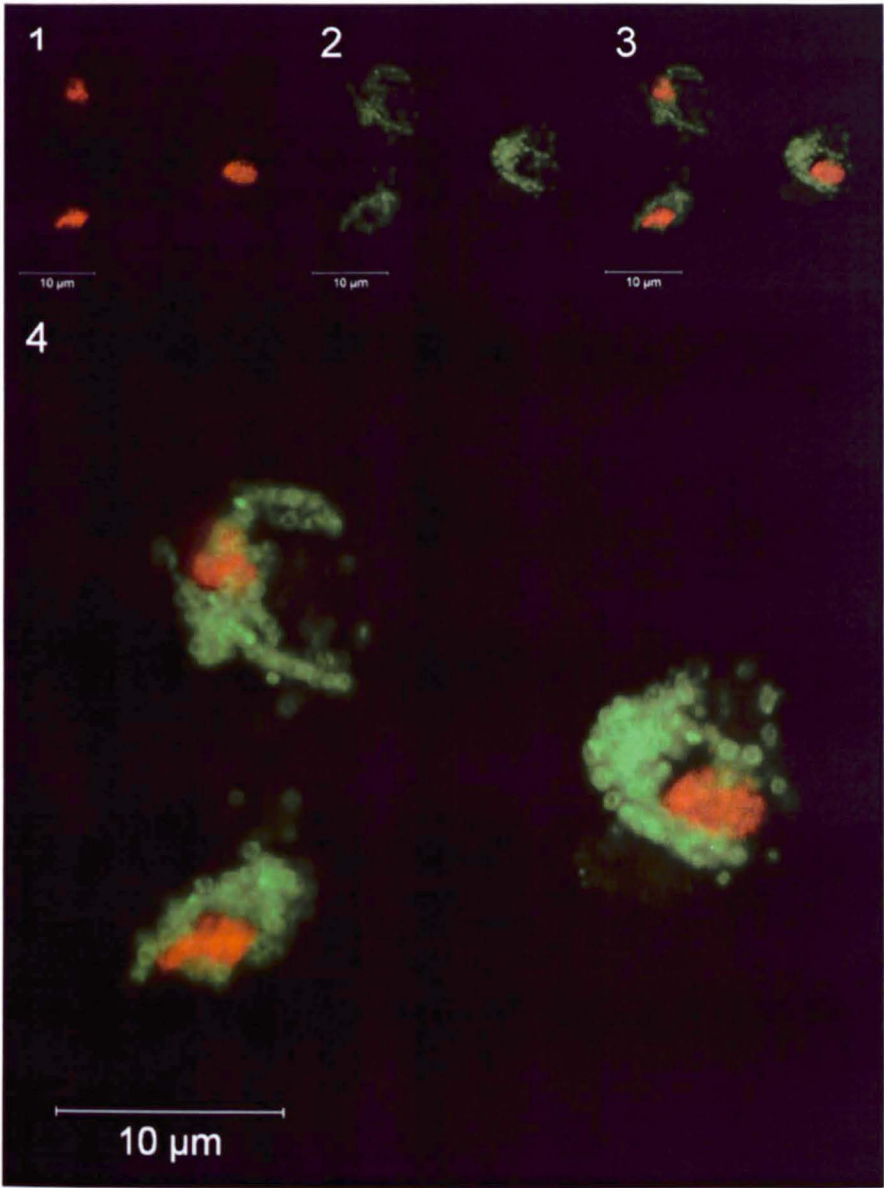


Figure 10.4.9 Intracellular distribution of Hup3c

Transformant cells expressing HA recombinant Hup3c were fixed and stained with nuclear marker PI (1) and anti-HA antibody (2). Combining these channels (3,4) reveal localization of the fluorescent tag to the membranous periphery of the hydrogenosome. Excitation of the PI channel was with a 561nm laser with light collected >600nm. The Alexa488 secondary antibody was excited using a 488nm laser with fluorescent light collected between 493-555nm. The image resolution was captured at 0.066 μm/pixel.

10.4.9 Appendix 22 Intracellular distribution of Hup3d

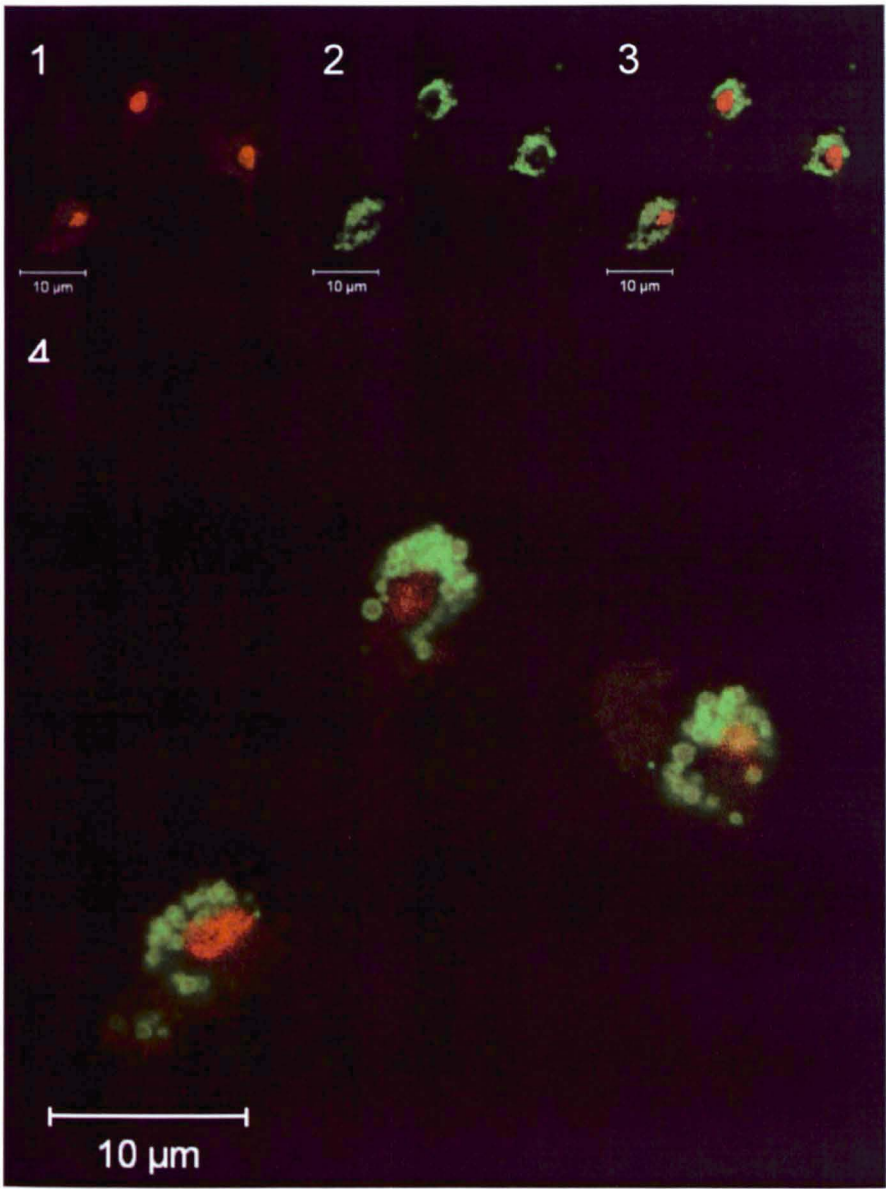


Figure 10.4.10 Intracellular distribution of Hup3d

Hup3d transformant cells were imaged and the distribution of PI (1) and immunodetected HA protein (2) were visualized in separate channels, recombination of these channels (3,4) show distribution of nuclei and hydrogenosomes in fixed cells. Excitation of the PI channel was with a 561nm laser with light collected >600nm. The Alexa488 secondary antibody was excited using a 488nm laser with fluorescent light collected between 493-555nm. The image resolution was captured at 0.088 μm/pixel.

10.4.10 Appendix 23 Intracellular distribution of LysoTracker bright bodies

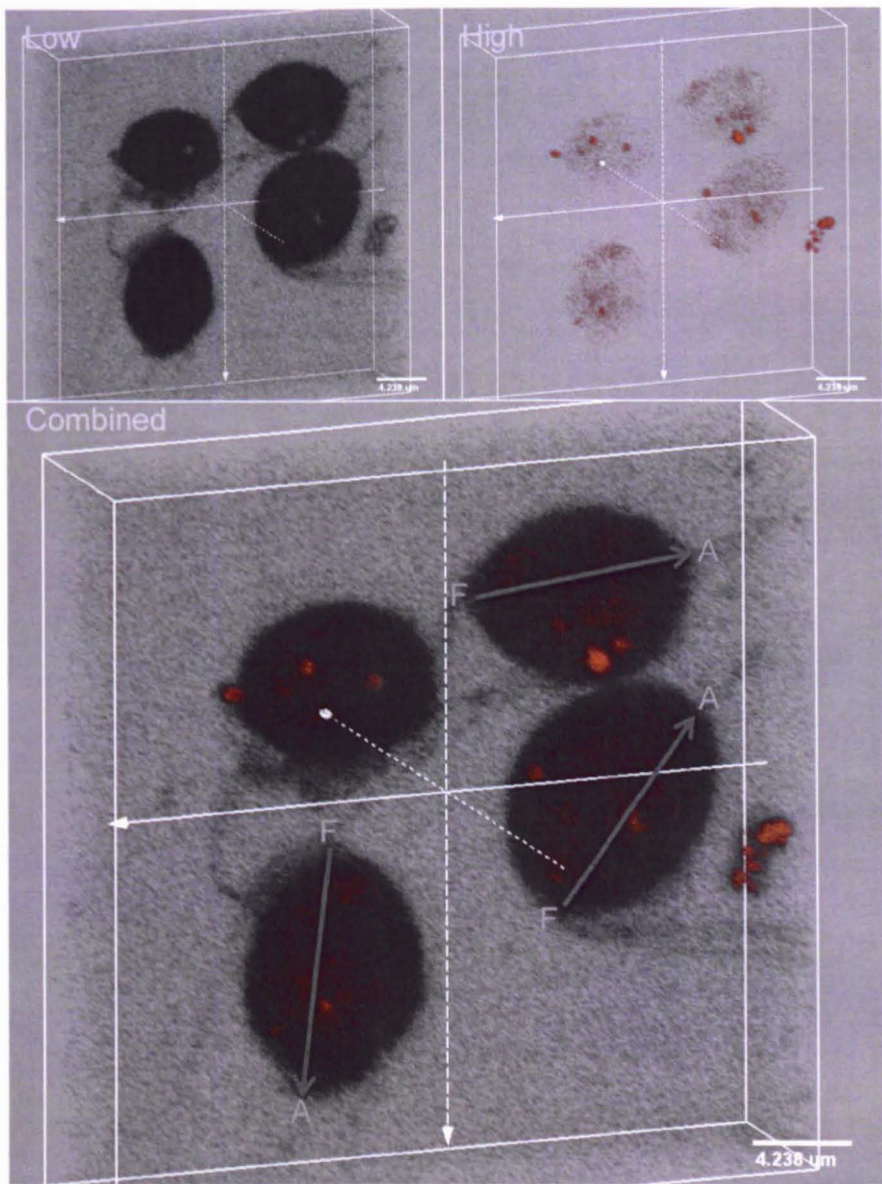


Figure 10.4.11 Localisation of LysoTracker bright bodies within 3D reconstructed cells

Using the bioView3D program 3D reconstructions of LysoTracker treated cells were reproduced. The programs alpha threshold was exploited in this figure to determine the localisation of the LysoTracker bright bodies. Using a Low threshold (Low), the noise generated by non specific LysoTracker binding was used to determine the structure and orientation of the imaged cells. A clear Flagella (F) – Axostyle (A) axis was determined for three of the cells, the fourth appears to have been imaged top-down through the fixed cell. Using a High threshold (High) Signal is reduced to only the more intense regions, which identify the bodies seen previously (Figure 6.3.12). Combined, these channels show that the distribution of these bodies is not related to the cell axis shown, nor do the bodies show any specific localisation patterns within the cell.

10.4.11 Appendix 24 measuring the dimensions of the hydrogenosome

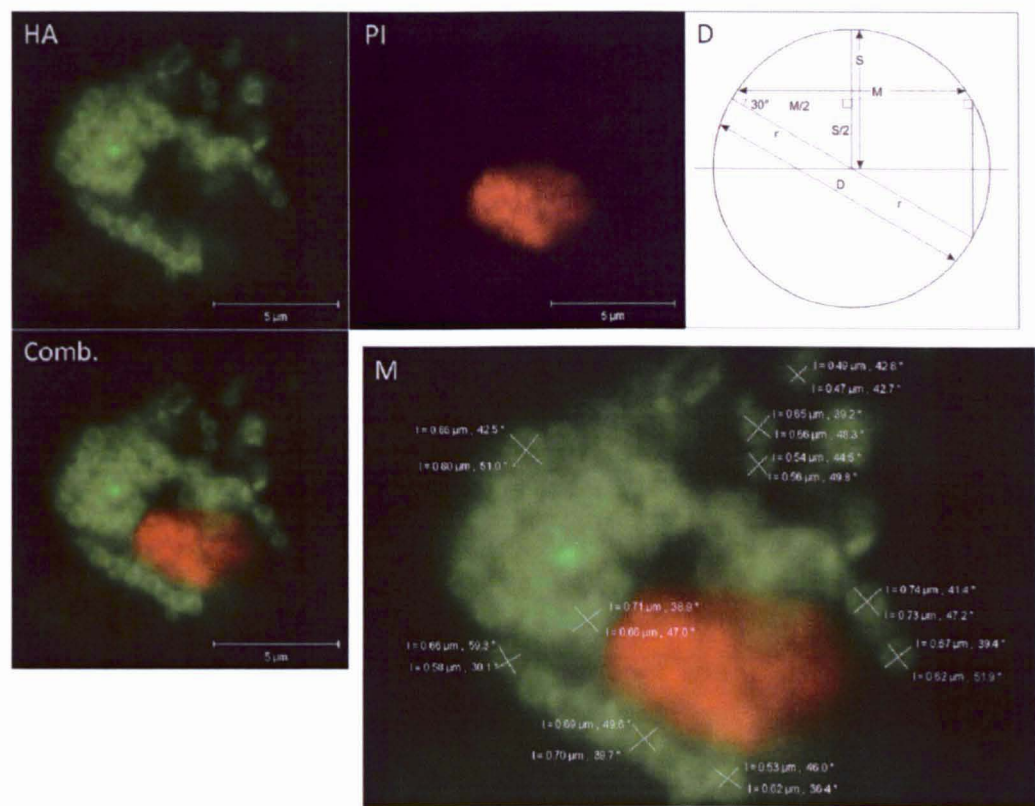


Figure 10.4.12 Microscopy analysis of a hydrogenosomal population in a Hup3 transformant

Whilst new microscopy methods might be employed to resolve the hydrogenosome in greater detail, the resolution offered by confocal microscopy does allow some analysis of the hydrogenosomal population. If each slice of the cell represents a 2D plane of organelles within that slice then each image reveals the diameter of organelles in that slice. Practically the Z-axis imaging performed in this study can not track organelles easily between slices, but the slices still reveal diameter information. These diameters can be measured, as shown above, albeit manually. Attempts were made to measure the organelles autonomously using MatLAB software, however the in-built image analysis software could not reliably discriminate between adjacent organelles. A collection of diameters can then be reasoned to determine the dimensions and diameters of a uniform spherical population (indicated above, the resolution obtained in microscopy is just sensitive enough to discriminate mean diameter from real diameter D), deviations from the expected distribution would inform about the size distribution of organelles- even if single organelles are not able to be tracked across planes. Although not shown, the distribution of diameters in all transformants are similar, though a larger sample population would be needed to confirm this.